



International Conference on Computational Science, ICCS 2017, 12-14 June 2017,  
Zurich, Switzerland  
Crowd Dynamics and Control in High-Volume Metro Rail  
Stations

Briane Paul V. Samson<sup>1</sup>, Crisanto R. Aldanese IV<sup>1</sup>, Deanne Moree C. Chan<sup>1</sup>,  
Jona Joyce S. San Pascual<sup>1</sup>, and Ma. Victoria Angelica P. Sido<sup>1</sup>

Center for Complexity and Emerging Technologies, De La Salle University, Manila, Philippines  
{briane.samson, crisanto\_aldanese, deanne\_chan, jona\_sanpascual,  
ma\_victoria\_sido}@dlsu.edu.ph

---

### Abstract

Overcrowding in mass rapid transit stations is a chronic issue affecting daily commute in Metro Manila, Philippines. As a high-capacity public transportation, the Metro Rail Transit has been operating at a level above its intended capacity of 350,000 passengers daily. Despite numerous efforts in implementing an effective crowd control scheme, it still falls short in containing the formation of crowds and long lines, thus affecting the amount of time before they can proceed to the platforms. A crowd dynamics model of commuters in one of the high-volume terminal stations, the Taft Ave station, was developed to help discover emergent behavior in crowd formation and assess infrastructure preparedness. The agent-based model uses static floor fields derived from the MRT3 live feed, and implements a number of social force models to optimize the path-finding of the commuter agents. Internal face validation, historical validation and parameter variability-sensitivity analysis were employed to validate the crowd dynamics model and assess different operational scenarios. It was determined that during peak hours, when the expected crowd inflow may reach up to 7,500 commuters, at least 11 ticket booths and 6 turnstiles should be open to have low turnaround times of commuters. For non-peak hours, at least 10 ticket booths and 5 turnstiles are needed to handle a crowd inflow reaching up to 5,000 commuters. In the current set-up, the usual number of ticket booths open in the MRT Taft Station is 11, and there are usually 6 turnstiles open. It was observed that as the crowd inside the station increases to 200-250 commuters, there is a significant increase in the increase rate of the turnaround times of the commuters, which signifies the point at which the service provided starts to degrade and when officials should start to intervene.

© 2017 The Authors. Published by Elsevier B.V.

Peer-review under responsibility of the scientific committee of the International Conference on Computational Science

*Keywords:* Crowd Dynamics, Crowd Control, Complex Systems, Mass Rapid Transit

---

## 1 Introduction

Designing and planning public spaces such as transport terminals, airports, malls, and stadiums are becoming more challenging nowadays because of the increasing need to support large volume

of people that exhibit different behavior and perform unconventional activities. It involves securing the safety of the people inside and near the area. In the Philippines, Metro Rail Transit has an average of 560,000 passengers daily. From 244 million passengers in 2004, to 415.5 million passengers in 2012, the passenger traffic in MRT stations increased by 71% in just a span of 8 years [2]. Aside from the increasing number of passengers, other contributing factors include its physical layout and various passenger behaviors. Thus, with the increasing instances of overcrowding in MRT3's terminal and high-volume stations, and interchanges, the Department of Transportation (DOTr) has been trying out different crowd control schemes to decrease overcrowding.

Crowd Dynamics is the study of how and where crowds form and move as the density of the crowd increases [10]. This involves modeling, simulating, and understanding the movement of individuals. Through knowing and predicting the movement of the commuters, how they use the space, and their activities in train stations, it will be easier to plan effective crowd control schemes, or floor layouts that could lessen crowd traffic. Currently, only an actual implementation of the plan can DOTC evaluate the effectiveness of a crowd control scheme. But through simulations, visualizing and understanding the crowd dynamics is easier since it offers a unique environment to test plans realistically by incorporating diverse traffic scenarios and human behavior [3]. Simulation results will lead to making informed design decisions. Although there are existing transportation models, there is a need for a model that considers the culture and environmental settings of the country.

This study aims to design and develop a crowd dynamics model of commuters in Metro Rail Transit or MRT3 stations that would help in evaluating their infrastructure preparedness. The agent-based model produced will be used to provide a tool that can aid in investigating the infrastructure preparedness of the MRT station. For the data gathering, the researchers developed a system called Swarm Plot, which allows the researchers and volunteers to annotate the data gathered from the MRT3 CCTV Live Feed. Swarm Plot produces floor field based from the annotations, and this floor field would be one of the factors that contributes to the behavior of the agents in the simulation system, SWARM-Simulation. Other factors that contribute to the behavior of the agents are the social forces adapted from [5] and some forces developed by the researchers.

The next section discusses some significant related works in creating crowd dynamics models. Section 3 details the data gathering method, as well as the annotation tool, SWARM Plot. Then, in Section 4, the architectural and detailed design of the SWARM simulation tool is discussed. Then, in Section 5, the results of the system validation and scenario analyses are explained. Lastly, Section 6 details a summary of the results and recommendations for future work.

## 2 Related Works

Many studies on crowd dynamics are currently focused on developed nations. One study [1] developed an agent-based model for public transportation which focuses on developing advanced driver assistance systems, semi- or fully autonomous buses, subway systems, and air transportation. Their agent-based model is able to capture the dominant socio-psychological factors operative within the presented Intelligent Transportation System context. This was inspired by the Metrobus system in Istanbul, Turkey. However, the study was not able to test emergent behavior in different simulation settings and compare the simulated data to real-world data.

Another study [7] developed a big model for transportation terminals which has individual models for rails, roads, general resources (human workers), pedestrians, trans-loading which is responsible for unloading and loading of freight between vehicles, vessels and storage, and lastly



Figure 1: CCTV live streams of the two entrances in MRT3 Taft Avenue station

models for infrastructures. It implemented and integrated their model using the simulation tool Villon. In their research, they were able to limit the knowledge of each sub-models about each other which made their model's interaction more effective.

[4] created a model that focuses on the group cohesion forces into certain pedestrian egress scenario. The scenario used in their simulation is a representation of a generalized pedestrian evacuation where individuals leave a nondescript venue through one exit. They applied a social force model [6]. They were able to conclude that the "tightness" of a group has an impact to the pedestrian egress and suggested that including goal dynamics is worth for the future study.

[9] also proposed an agent-based approach for crowd dynamics simulation. But in their approach, they involved a multi-level architecture wherein the social force model was applied as the low level component and the agent's high-level component is what makes them capable of navigating, approaching the destination and interacting with others. Their simulation setting was based on the Xi'an Railway Station in China, and was developed in C++ and Qt Creator. One of the contributions of their work is that their model involves architecture with higher levels, such as path planning, and each part of their model is extensible because it is modularized. The commuters in their model are assumed to be familiar with the situations in the station, which is a limitation to this model as the familiarity of the commuters to the station could affect the path planning level.

### 3 Data

This section will discuss the different data sources, gathering methodology, and data processing.

#### 3.1 Crowd Formation

In gathering data for crowd formation, the live CCTV feeds of the MRT3 Taft Avenue station's concourse area was accessed to observe the ticketing booths and waiting areas in real time.

Figure 1 shows a screenshot of the two entrances in MRT3 Taft Avenue station. On the left is the entrance mainly for stored value ticket holders and those with special needs. The right image shows a view of the ticketing booths and the queues of people buying mainly single journey tickets. These two views were the main sources of crowd formation data for the study.

An image scraper was developed to automatically acquire images from the CCTV live feeds. Since we only want to focus on where passengers usually wait, queue, and pass through, the tool only saves an image every 30 seconds. The interval is enough to track where crowds usually form. The data gathering lasted for a week and images were scraped from 5AM to 10PM, covering the normal operating hours of the station. The captured images have a dimension of 640x480 and a total of 728 live feed images were captured.

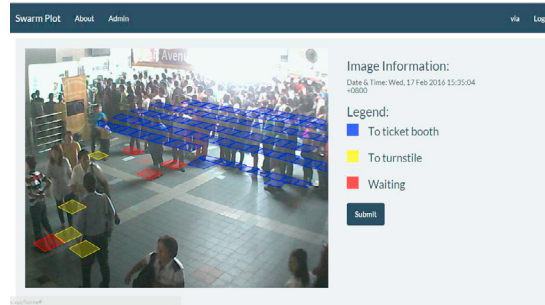


Figure 2: Annotating the crowds forming at the entrance for single journey passengers

### 3.2 Image Annotation

After automatically gathering images from the MRT3 live feeds, they were annotated with the location of the people and the activity they are most likely doing. A stand-alone web application was developed to help in annotating the images.

Since the images from both live feeds were not taken overhead, an image mapping tool [8] was used to get the coordinates of each cell on each entrance. The size of each cell is 2x2 of the floor tiles which was based on the empirical observation of the researchers during their on-site data gathering. The coordinates produced from the image mapping tool was integrated in the Swarm Plot web application through CSS.

The researchers recruited students who also ride the MRT3 as annotators. During annotation, Swarm Plot randomly serves images from any of the two live feeds (See Figure 2). Annotators then select the cells where commuters are and tags whether they are walking towards the ticket booth, to the turnstile, to the exit or is currently waiting, by clicking the cell a corresponding number of times - click once if the commuter is going to a ticket booth, click twice if the commuter is going to a turnstile, click thrice if the commuter seems to be standing still or waiting, and click four times if the commuter is going to an exit. Each image can only be annotated by three unique annotators.

After all annotators were done, Swarm Plot sifts through each recorded annotation and determines the frequency each cell is clicked for a specific action. Four separate floor fields were generated, one for every action that the web application supports (1) walking towards a ticket booth, (2) walking towards a turnstile, (3) idly waiting, and (4) walking towards an exit. Two versions of each static floor field were generated, one representing peak hours and another for non-peak hours, for a total of 8 static floor fields. These will be used later by the simulation.

### 3.3 Passenger Distribution

Data for the passenger volume for the month of July 2014 has been provided by the Department of Transportation. For each origin-destination pair of stations, there is data on the total number people that used the station based on the type of ticket, either SJ (Single Journey) or SV (Stored Value). Aside from passenger volume, the data set also contains information on the number of passengers entering and exiting the stations per hour.

## 4 Architectural Design

The architecture has three main modules, namely the image scraper, Swarm Plot, and Swarm Simulation.

First, the image scraper deployed onto a server instance collects images from the public MRT3 Live Feed every 30 seconds, and stores them locally onto the instances SSD.

After the scraper finishes collecting the images needed, they are downloaded onto the researchers' development machines via FTP, and are uploaded onto a system developed for image annotation called Swarm Plot. Swarm Plot allows users to annotate cells which have commuter agents performing specific actions.

Once all the images are uploaded onto Swarm Plot, they are served to users such as the researchers themselves, and volunteer annotators. These users will manually annotate each image, specifying on which cells there are commuters performing different types of actions on. These annotations will be used in order to create a static floor field.

After all the images have been annotated, the researchers generate floor fields from Swarm Plot based on time of day and action being done by the commuter. These floor fields will serve as one of the factors in the to-be simulation's input, and additionally, will serve as the basis of what the simulation's layout will be.

The simulation itself takes multiple types of input - the floor field that was generated from Swarm Plot earlier, simulation parameters initially set by the user, and data collected from various sources such as the hourly passenger rate and the passenger distribution. The simulation uses a commuter generator which handles the production of the commuter agents for every time step. The number of commuters that will be spawned in the simulation will be based on the passenger distribution from the data gathered.

The virtual concourse area, comprised of the guard and the barriers are generated based on the layout initially set from the floor field. In addition to this, the simulation also uses the passenger rate and passenger distribution to determine how the simulator will be generating commuter agents.

The agents that will be inside the virtual concourse area are the commuter agents, turnstiles, and ticket booths. The commuter agents interact with the ticket booths if they need to purchase tickets, and interact with the turnstiles if they already have a ticket.

After the simulation has completed, it uses all the data gathered from the commuters, ticket booths, and turnstiles in order to generate reports that would detail what had transpired during the time the simulation was running. The reports may help in analyzing the infrastructure preparedness of the MRT3. It calculates relevant data such as the average time of a commuter agent in the station, the crowd density and crowd flow. The reports generation features station analysis which keeps track of the crowd density and the number of commuters the turnstiles and ticket booths can handle for a specific time period; and travel time analysis which keeps track of the time it takes for a commuter agent to go from one point to another.

## 5 Detailed Design

The web application built for the simulation serves data such as the annotations, distributions, and other pertinent data needed for the simulation from its PostgreSQL database to the simulation powered by AgentScript, an agent-based modeling framework. The simulation input required are the floor field from Swarm Plot which is used to set the map layout and determine the frequency each cell is stepped on by a commuter, the frequency of which each turnstile is used to enter the platform area which is used to aid the commuter's decision on which turnstile

to use, the passenger distribution data that determines the destination of commuters which aids the commuter's decision on which ticket booth to line up for (since there are ticket booths that only serve tickets for specific destinations), and the distribution that determines the number of commuters that go into the station every hour which aids the simulation in determining the frequency and number of commuters to generate.

The number of commuter agents entering the simulation is dictated by the passenger distribution. This number is spread out randomly across 3,600 seconds, which the simulation uses to determine the number of commuter agents to spawn for every second. Each attribute as detailed in the previous section is also determined randomly based on the data collected by the researchers. The researchers implemented a crowd control system similar to that of the Taft MRT station. This module is activated whenever the lines for the ticket booth are at full capacity. The entrance where single journey commuters usually enter delays the entry of commuters, and continues letting commuters in only when the ticket booth lines are reduced to half of its capacity.

## 5.1 Agents

In the simulation, the commuter agent is initialized by randomly chooses its destination and whether it uses a Single Journey (SJ) or Stored Value (SV) ticket. After which, based on whether it uses an SV or SJ ticket, it randomly chooses a turnstile or ticket booth to be its target goal, respectively. It chooses its target goal based on the length of the queue for the station component. The commuter agent favors the station component with the shortest queue.

After which, it is spawned into the virtual concourse area. This refers to the open space that Commuter Agents can occupy within the simulation. This area is enclosed by barriers that are set from the initial floor layout.

Aside from the commuter agent, the ticket booth agent serves single journey commuter agents tickets and act as the transitional state between going to the ticket booth and going to the turnstile, while the turnstile agents serve all types of commuter agents and lets them exit the simulation.

## 5.2 Social Force

[6] introduced the social force model stating that the behavioral changes of pedestrians are guided by social fields and forces that drive a pedestrian to move as a reaction to the information he perceived in the environment. The simulation uses a modified version of social force path-finding method that takes into account both commuter agents interactions with other commuter agents in combination with the static floor field. The model yields a vector force indicating the direction of the agent's next movement. After getting the cell representing the agent's next position, its cell value is compared to the value of its adjacent cells and picks the cell with the highest floor field value.

## 6 Results and Analysis

Different strategies were employed in evaluating the developed crowd dynamics model and the different simulation scenarios and assessing the station's infrastructure preparedness. One method used for validation was the historical data validation to verify whether the data input in the simulation is used by the system and matches with the system output. This involves comparing the percentage distributions generated by the model with that of data gathered

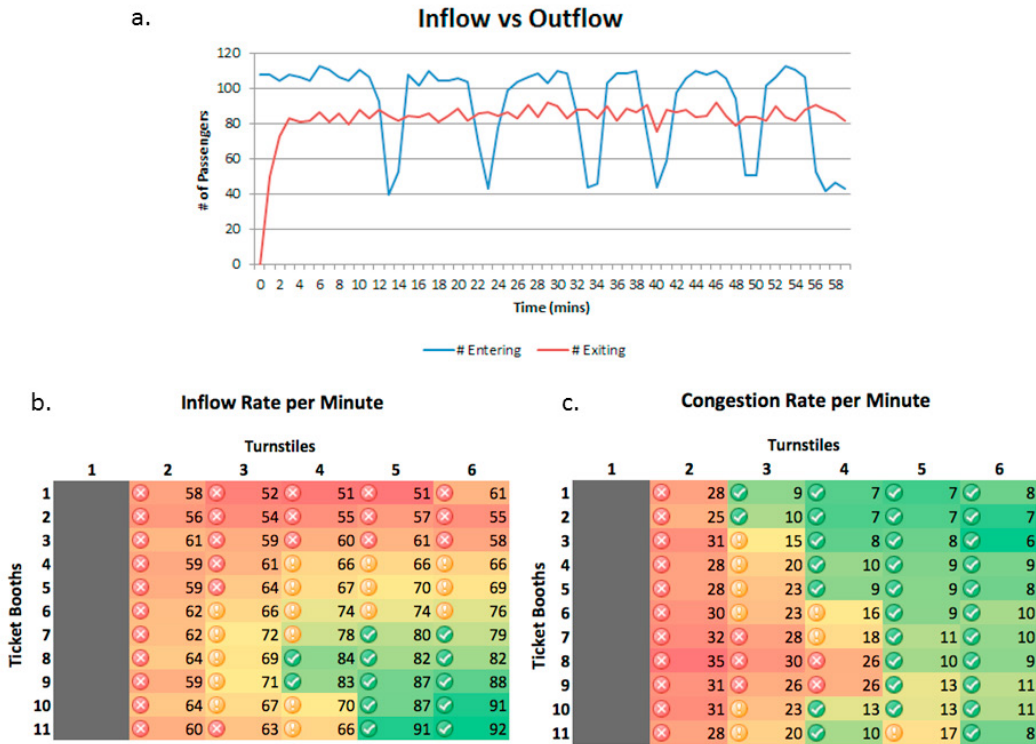


Figure 3: a. Commuter Inflow vs Outflow, b. Inflow per minute, and c. Congestion per minute

from Taft MRT station. The researchers ran three runs of simulation and took note of the floor fields, passenger traffic, passenger type distributions and passenger turnaround times. Agent state checking was also done to verify if the commuter agent passed through all the required states.

The researchers also performed Parameter Variability - Sensitivity Analysis to determine the effect of the number of enabled ticket booths, and turnstiles to the travel time and crowd congestion of commuters inside. The results of these tests were also used for the analysis of the infrastructure preparedness. For this analysis, the station congestion, average turnaround time and the crowd congestion on passenger turnaround time were observed.

### 6.1 Station Congestion

To observe the station congestion, the inflow vs outflow graphs produced by Swarm Simulation were used. Inflow vs Outflow graph (Figure 3a) refers to the number of passengers entering the station in relation to the number of passengers passing through the turnstiles. This is used to assess if the number of turnstiles enabled can handle the rate of entering passengers. The researchers computed for the inflow rate of commuters per minute (Figure 3b), as well as the congestion rate per minute (Figure 3c). The congestion rate is computed by subtracting the outflow rate from the inflow rate. After running it was found out that it is only when at

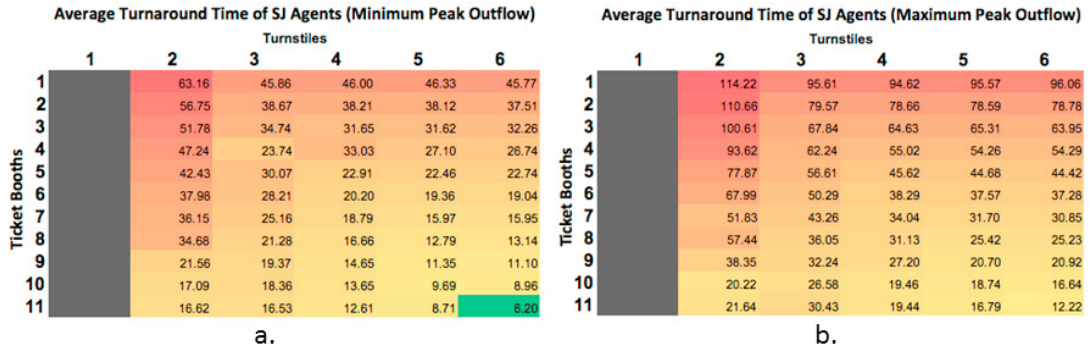


Figure 4: a. Average Turnaround Time of SJ commuters: minimum commuter flow of 3,696 commuters, and b. Average Turnaround Time of SJ commuters: maximum commuter flow of 7,386 commuters

least 7 ticket booths and 5 turnstiles are opened that outflow rate can handle a high inflow of commuters while minimizing overcrowding inside the station.

### 6.2 Average Turnaround Time

Average turnaround time refers to the time taken by commuters to go from entrance to turnstiles. The researchers derived the minimum and maximum number of passengers during peak and non-peak hours from the DOTC-MRT Passenger Traffic data, and ran all the 66 possible settings of the ticket booths and turnstiles. Aside from assessing the effect of the number of ticket booth and turnstiles, the researchers also want to find out which configuration of the ticket booth and turnstiles will be able to handle the number of passengers inside the station at any given time. Single journey commuters find it acceptable to have a turnaround time of 6 minutes during peak hours and 4 minutes during non-peak hours. For stored valued ticket commuters, they expect to go past the turnstiles after 1 minute. This shows the quality of service of the station, where lower turnaround times mean better service. It is also used to determine the minimum station configuration without significantly degrading the service.

When the current implementation during peak hours in Taft MRT station accommodates 3,696 commuters (Figure 4a), it was found that at least 11 ticket booths and 6 turnstiles should be open to accommodate the expected commuter flow without significantly degrading the service for both Single Journey commuters and Stored Value commuters. However, when the station accommodates 7,386 commuters (Figure 4b), it needs more than 11 ticket booths and 6 turnstiles to provide quality service to the commuters. Even though all ticket booths and turnstiles are open, the average turnaround time of SJ commuters still reaches as 12 minutes. As for non-peak hours, at least 10 ticket booths and 5 turnstiles should be open.

### 6.3 Crowd Congestion on Turnaround Time

Crowd congestion on turnaround time refers to the time taken by commuters to exit the simulation in relation to the number of commuters inside the station. This would show the effect of crowd congestion inside the station to the turnaround time of the commuter agents. The researchers also made use of the scatter plot produced by the Swarm Simulation to plot the average turnaround time of the agents for each distinct number of passengers. For example,



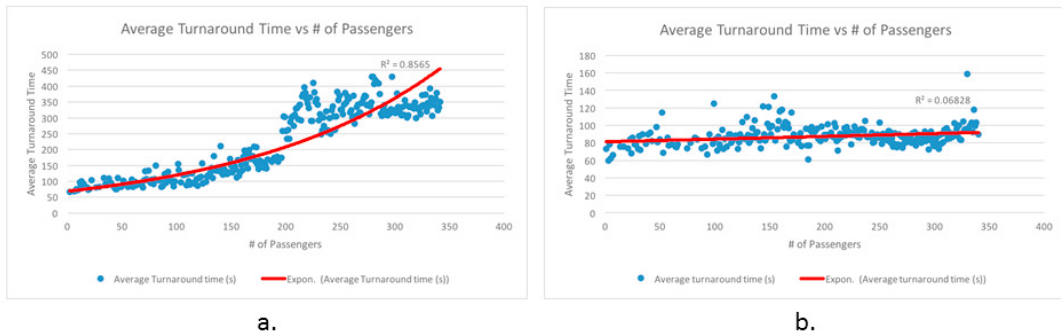


Figure 5: a. Crowd Congestion on Turnaround Time for SJ commuters (11 ticket booths, 6 turnstiles enabled), and b. Crowd Congestion on Passenger Turnaround Time for SV commuters (11 ticket booths, 6 turnstiles enabled)

during the whole run of the simulation, there were 100 agents that was spawned inside while there were 400 passengers inside, the average turnaround time of these 100 agents is then plotted to the graph.

For the graphs obtained for SJ commuters, the elbow of the graph was observed. Elbow refers to the point of the graph where there is a sudden huge increase in the increase rate of the turnaround time. The elbow shows how many passengers can the MRT station handle before it degrades its service and the passengers are having difficulty in moving within the station. It was found that as more turnstiles are enabled in the simulation, the elbow decreases and as more ticket booths are enabled in the simulation, the elbow increases. However, the elbow of the graphs is relative to the maximum number of passengers inside and their average turnaround time. In the current implementation at the station wherein 11 ticket booths and 6 turnstiles are enabled (Figure 5a), it was determined that as the crowd inside the station increases to 200-250 commuters, there is a significant increase in the turnaround times of the commuters. It is at this point that MRT3 officials should start to intervene because the services of the MRT3 station starts to degrade.

As for SV commuters, it was found that when less than 3 turnstiles are enabled (Figure 5b), the turnaround times of SV commuters reach as high as 1,500 ticks (25 minutes). The lesser the enabled turnstiles in the simulation, the number of commuters inside the station increases faster and the turnaround time also increases. The turnaround time is almost constant in some scenarios, especially when there are more turnstiles enabled. This is because more turnstiles means less crowding inside the station and the SV passengers are able to exit the turnstiles without much delay. Therefore, the turnaround time of SV commuters are dependent on the number of turnstiles enabled and the number of ticket booths open does not have a significant effect on their turnaround times. With the current set-up of the MRT3, there is no elbow for the turnaround time of SV commuters. Their average turnaround time is 1-2 minutes only.

## 7 Conclusion

This research designed and implemented a simulation tool that would help in understanding the crowd behavior in the concourse area of the MRT3 Taft Avenue station. It was designed to help in planning strategies to lessen overcrowding. The architecture has three main modules: Image Capturing, Swarm Plot, and Swarm Simulation.

The model allowed the simulation of various scenarios such as disabling ticket booths, disabling turnstiles, disabling both ticket booths and turnstiles, adding ticket booth delays, adding turnstile delays, and restricting the stations being served in ticket booths. Results showed that at least 7 ticket booths and 5 turnstiles are needed to minimize overcrowding. For peak hours, at least 11 ticket booths and 6 turnstiles are needed to handle the expected crowd inflow, which may range from around 4,000 agents to 7,500 agents. As for non-peak hours, at least 10 ticket booths and 5 turnstiles are needed to handle the expected crowd inflow ranging from around 100 to 5,000 agents. Based on the results, it is concluded that the MRT3 Taft Avenue station could only handle up to 200-250 passengers at the same time inside the station. Letting more people in would start to degrade the service of the station.

There are still some aspects of the model that were not considered by the researchers because there is no available data yet. An example would be the different types of passengers and their behaviors which may contribute to the crowd dynamics of the model. In addition to this, the data that the researchers were able to gather are from 2014, this means that there could be a lot of changes already in the passenger traffic reports for 2015 and 2016. Another recommendation is to conduct surveys to the MRT3 passengers regarding the turnaround time, and the factors that affect their decisions on ticket booth and turnstiles.

To become a more effective tool for planning, it is recommended that components' positioning and their quantity within the simulation can be dynamically set by the user. Through this, users can analyze the effect of adding more ticket booths and turnstiles if they could install more, and the effect of the positioning of the components to the crowd behavior. It is also recommended to have an easier way to change the map layout, or floor layout of the station.

## References

- [1] E. Adamey, K. Arda, and U. Ozguner. Agent-based passenger modeling for intelligent public transportation. In *Intelligent Transportation Systems - (ITSC), 2013 16th International IEEE Conference on*, October 2013.
- [2] Philippine Statistics Authority. Rail, water, land and air transportation statistics 2003-2012, 2014.
- [3] John R. Clymer. *Simulation-Based Engineering of Complex Systems*. Wiley-Interscience, New York, NY, USA, 2nd edition, 2009.
- [4] Andrew Collins, Terra Elzie, Erika Frydenlund, and R. Michael Robinson. Do groups matter? an agent-based modeling approach to pedestrian egress. *Transportation Research Procedia*, 2:430 – 435, 2014.
- [5] Dirk Helbing, Illes Farkas, and Tams Vicsek. Simulating dynamical features of escape panic. *Letters to Nature*, 2000.
- [6] Dirk Helbing and Pter Molnar. Social force model for pedestrian dynamics. *Physical Review E*, 1995.
- [7] Michal Kocifaj, Michal Varga, and Norbert Adamko. Modelling interactions in agent based models of transportation terminals. In *Digital Technologies (DT), 2014 10th International Conference on*, Zilina, Slovakia, July 2014.
- [8] Adam Maschek. Online image map editor.
- [9] Qiu Qin and Junhu Wei. An agent-based approach for crowd dynamics simulation. In *Intelligent Computing and Intelligent Systems (ICIS), 2010 IEEE International Conference on*, Xiamen, China, October 2010.
- [10] G. Keith Still. *Crowd Dynamics*. PhD thesis, Department of Mathematics, University of Warwick, 2000.