Evaluating Aspects of WeChat Use for Information Sharing During a Campus Attack Event Using Agent-Based Simulation

Steven Beattie*, Jing Yang (Sunny) Xi and Wai Kin (Victor) Chan

1 Department of Data Science and Information Technology, Tsinghua-Berkeley Shenzhen Institute, Tsinghua University, Shenzhen, 581055, P.R. China
2 Department of Environmental Science and New Energy Technology, Tsinghua-Berkeley Shenzhen Institute, Tsinghua University, Shenzhen, 581055, P.R. China
Email: bstevie@umich.edu; sunnyx@berkeley.edu; chanw@sz.tsinghua.edu.cn

Abstract. Information sharing during emergency events often takes place through popular social messaging apps. An emergency response area of increasing concern in recent times is that of mass attacks on school campuses yet work studying information sharing systems in these situations is scarce. In this exploratory work, we used data from various sources to construct an agent-based simulation of a hypothetical knife attack on a university campus in southern China. We model the information sharing system after WeChat and evaluate the impact of group and individual messages on attack response and outcome. Evaluation metrics such as how students become alerted to the attack, the maintenance of accurate information, and casualty levels, are tracked. Results suggest that group messages may be most helpful despite being perceived as less trustworthy, and that the presence and structure of information sharing systems has a significant effect on response to campus attack events. Future work in the area is suggested.

Keywords. Emergency response; agent-based; social network; information sharing; WeChat; campus attack.

1. Introduction

Information sharing dynamics will undoubtedly have an effect on a community’s response to an emergency situation; they may also have a significant impact on the outcome of such a situation. An emergency response area of increasing concern in recent times is that of mass attacks on school campuses. Despite its relevance, this area remains an under-explored area of emergency response work with respect to the dynamics and effects of information sharing. With this situation in mind, we used data from various sources to construct an agent-based simulation of a hypothetical knife attack on a university campus in southern China.

Within this simulation, we have modeled the information-sharing and security alert systems currently in use on the campus: WeChat and emergency calling. This combination constitutes the de-facto response infrastructure on many Chinese campuses, and internationally the situation is largely the same despite using other messaging or social media platforms than WeChat. As an initial step in this research direction, we study the effects of the ability to send messages to friends and/or groups on various situational factors, for example the level of community awareness, the maintenance of accurate information, and how quickly the attack is ended.
Students today are more connected electronically to the world around them than ever before. They routinely use platforms such as Facebook, Twitter, and WeChat to gather and share information on what is happening in the world around them. Currently, these platforms constitute the quickest and most direct routes of information sharing in an on-campus emergency situation, despite not necessarily being designed for this purpose. The better that we can understand the information sharing and update dynamics in the systems currently used, and their effects on situation outcome, the more effective future efforts to develop better emergency information sharing systems can be.

2. Related Work

Much work has been done looking at the effects of Twitter use in emergency response scenarios. Abedin and Babar analyzed Twitter’s role in the response to the 2014 Australian bushfires [1]. Hughes and Palen have studied the use of Twitter during emergency events and suggest that Twitter can be adopted in “emergency management... as a way of getting information to the public” [2]. However, not all information online is true; for example, Thomson et al. studied tweets from different sources during the Fukushima Disaster and ranked them on their trustworthiness [3]. Around 70% of information tweeted by the public are based on credible sources. In the case of campus attacks specifically, many general event analyses exist – including a detailed thesis by Linger which analyzed 186 of the most recent mass shooting events in the U.S. There also exist many qualitative analyses of information sharing in emergency situations generally [4].

Agent-based simulation is a popular approach to researching various aspects of emergency response, with some looking at the case of campus attacks specifically. For example, Zhang et al. looked at evacuation during violent attacks in open public space [5]; Ma et al. considered response to sarin gas attack in a railway station [6]. Xi and Chan, as well as Briggs and Kennedy have explored the risks and benefits of fighting back during an attack [7, 8]. Others including Anklam et al. have considered the possibility of encouraging “concealed carry” of firearms on campuses [9]. There also exists multiple agent-based evacuation models; some of these incorporate psychological models of agent behavior (e.g. Belief-Desire-Intention by Okaya and Takahashi, but do not seem to directly consider information-sharing dynamics [10].

While related work has been done, it seems that there may not be previous work specifically modeling and evaluating the effect of information sharing systems and behaviors in a campus attack scenario.

The agent-based simulation for this project was coded in NetLogo. The model itself was initially developed using the code produced by Xi and Chan in modeling a campus attack scenario without considering communication dynamics [7]. In the current project over 90% of the code is independently developed, but some of the original code remains in the current model. This code mostly relates to environment creation – for example, overlaying a node network onto a campus map and implementing the links between nodes to facilitate agent movement.

3. Methodology

3.1. Geographical Information System (GIS) Mapping

Figure 1 shows the constructed environment for the NetLogo simulation model. The environment is a GIS map of the campus studied, overlaid with 636 nodes and 720 links.

3.2. Campus Survey

In a voluntary-response survey, we asked 165 students on campus about their social circles and knowledge of dealing with emergencies on campus. The results (shown in table 1) helped inform model development and parameter tuning.
3.3. Agent-Based Model

The simulation model begins with 2120 students, one attacker, and two police. In accordance with survey results, only 20% of students can make a police report. All students have the option of sending text messages (if the test scenario allows it), but to keep this initial work simpler only 4 total messages can be sent by each student. When the attack begins, students who are aware of the attack must choose how to respond. Attack awareness can result from a student being at the attack location, or from receiving a message about the attack. Students adopt a common “run, hide, fight, then call police” strategy, and may also send a message or file a police report [11].

As the testing scenario allows, messages can be sent to a group chat or directly to a friend. Messages contain information about the current attack location; they are categorized as either correct, incorrect, or incomplete. Based on the source of the message and the context in which the message was sent (i.e. whether the sender was on the then-current attack location or not), the message will be believed with various probabilities.

In the case of an incomplete message, students will try to interpret the message. Naturally, there are three possibilities for interpretation: the student interprets and considers correct or incorrect information, or discards the message if no sense can be made of it. Attack location information is categorized as correct or incorrect based on how close the location is to the current attack location; as the situation changes (i.e. as the attacker moves), students’ accurate information about the situation quickly becomes incorrect if they are not able to update their perception of the situation appropriately.

Through report calls, students can alert police officers to the situation. This requires time to be spent communicating pertinent information (e.g. the reporter’s believed attack location), conveying that information to personnel, and planning a response. Once a response is initiated, police officers move to engage the attacker using the most up-to-date information that they are provided with. When the attacker is found and engaged, the simulation ends.
4. Discussion of Model Parameters
In a second, less structured survey, students indicated that their hypothesized action choices were made effectively at random, but that a majority of students would first consider security reporting before considering alerting friends or a group. Accordingly, equal probability exists for each possible choice in a given situation. Students are restricted from making a security report while at the attack location, reflecting the standard strategy of “run, hide, fight, then call for help” [11]. Students may have varying trust levels in authority; thus, we use a 50% chance that a student will make a report when possible; otherwise, they will choose to alert friends before alerting the group (when possible).

When evaluating new messages, students will consider both the source and the content. Complete messages sent from the attack location will presumably somehow indicate this fact; this “logos” appeal will presumably not be as strong in messages sent from elsewhere. Also, messages will inevitably carry some level of “ethos” appeal. We took a 50% probability to indicate a lack of conviction one way or another about message content, and 100% probability as the willingness to adopt information from a source without question. We set (and tuned) the probabilities of message belief within this range, in order of stronger appeal: central message → friend message → group message. This ordering also reflects the empirical evidence seen in various works, including Okaya and Takahashi [10]. However, there are no central messages included in the current tests. This reflects current circumstances at the university studied.

For the following parameters, we set values that produce a possible “worst case” scenario yielding high average casualty levels. With these values, the attacker exhibits reasonable and active movement patterns, and students are continually challenged in updating to currently-accurate information about the situation. This allows us to better study the effects of inaccurate or incomplete information generation in current systems. We set the probability of sending an incomplete message to 33% (data on appropriately similar circumstances is scarce; this value is in loose reference to Thompson et al.) [3]. A maximum knife attack rate of once per 5 seconds was set; time to produce a knife fatality is highly variable, so in our project we consider a casualty as at least injured. Three attack locations on campus were chosen, the attacker has only one minute (minimum) at each attack location, and has a “swiftness” factor of 15% (the probability that, if the attacker is not at an attack destination, they will continue towards one even if there are students at the current location).

We also consider the time it takes to submit a report and initiate a security response. We set the report submission time as uniformly random between 1 and 3 minutes, and set the report processing time as uniformly random between 1 and 2 minutes, giving an overall time to initiate a security response as uniformly random between 2 and 5 minutes. This reflects the findings of Linger and Anklam et al. [4, 9]. Due to a lack of empirical data, we set the probability that a student becomes aware of (or updates to the correct belief about) an attack happening at their location as 50%; through testing, we verified that this parameter value does not exert a significant primary effect on our variables of interest. Finally, we should note that each tick (time step) in our simulation is the equivalent of 5 seconds in real time.

5. Experiments
In this exploratory work, we wanted to begin studying the impact that individual (friend) and group messages has on response and outcome in a campus attack scenario. In the evaluation of impact, we considered 13 metrics: casualties (figure 2), number (and rate) of students alerted (figure 3), percent of alerted students with correct info (figure 4), whether students’ alert was in person or from a message (figures 5 and 6), total number of messages sent (figure 7), number of messages sent to group or friend specifically (figures 8 and 9), the rates of correct and incorrect information update associated with both (figure 10-13). Table 2 presents average end-of-run results for pertinent metrics from each scenario.
Figure 2. Total casualties.

Figure 3. Number of students alerted.

Figure 4. % of alerted students with correct info.

Figure 5. Number of students alerted by message.

Figure 6. Number of students alerted in person.

Figure 7. Total number of messages sent.

Figure 8. Number of messages sent to group.

Figure 9. Number of messages sent to friends.
For our experiments, we considered four scenarios: Scenario 1, the baseline scenario with no messaging; Scenario 2, which allows for only friend messages; Scenario 3 for only group messages; and Scenario 4 with both group and friend messages allowed. We collected data from each time step of 2265 simulation runs, and collected end-of-run measurements as well. Of these runs, 540 were from Scenario 1; 575 were from Scenario 2; 575 were from Scenario 3; and 575 were from Scenario 4. The data were cleaned and separated into run info and end-of-run info. Each run and its corresponding end-of-run info were separated into lists corresponding to the scenario simulated. From each scenario, an average end-of-run report was compiled. Then, a dataset representing an average run from each scenario was compiled. Each dataset was designed to be of average length for a run from the respective scenario, and each run with entries for a sampled tick were included. With datasets constructed for each scenario, we proceeded with analysis.

6. Discussion of Results

6.1. Runs with No Messages (Scenario 1)

In Scenario 1, we saw that the number of students alerted was comparatively low (figure 3). This scenario is very far away from being an accurate representation of the event as it unfolds in real life; our purpose in including results for this scenario is that it illustrates the potential benefit of incorporating information-sharing dynamics into a model. Info sharing dynamics may not be helpful for every application, but these results illustrate how considering these dynamics could help build a more complete understanding of these kinds of emergency events.

In future work, the number of casualties alerted at the time of their attack should be tracked; we believe that this would be an especially interesting dimension to consider in the case of no-messages. A notable result in Scenario 1 is the fact that students were able to maintain correct information so effectively. In this scenario, the only way to update to a correct belief is to be continually (or...
repeatedly) at the attack location; it would be interesting to see if this trend is altered by updates to students’ evacuation route planning based on new info. (Currently, route update is not in the model).

### Table 2. End-of-run results metrics.

| Scenario       | Evaluation Metrics (Scenario Averages) | Simulation Time (ticks) | % w/ Correct Info | Num Alerted | Total Casualties |
|----------------|----------------------------------------|-------------------------|-------------------|-------------|------------------|
| 1 (None)       |                                        | 120.80                  | 66.77%            | 779.06      | 92.32            |
| 2 (Friend)     |                                        | 119.10                  | 67.41%            | 1271.24     | 90.39            |
| 3 (Group)      |                                        | 111.34                  | 73.78%            | 1828.66     | 82.70            |
| 4 (Friend & Group) |                                      | 111.08                  | 73.21%            | 1827.85     | 82.23            |

6.2. Group Messages (Scenarios 3 and 4)

We can see a clear difference between scenarios in the length of time the attack lasts, and can see a strong correlation between the total attack time and the number of casualties incurred (table 2). Thus we consider that casualty levels are a function of attack length. How long the attack lasts is a function of how quickly a report is made, or how quickly a student decides to fight back; both of these depend on how many students are alerted, and how quickly.

Group messages seem to be far more effective at alerting the student community than private messages, which seems an intuitive result. But individually they are significantly less trusted (25% less in this model) than other message types. This skepticism leads to a far lower rate of belief update from group messages, and individual messages can be seen as less effective. Also, the possibility for misinterpretation is generally higher for group than friend messages (33.3% vs 25%) which can explain why the number of students adopting incorrect information from group messages rises steadily as number sent increases.

At the same time, reliance on group messaging did not seem to negatively affect the overall ability of students to maintain current and accurate understanding of the situation (figure 4), and lower levels of incorrect info are seen (figure 14). This is an example of “self-correcting” or “collaborative filtering” phenomena seen in social networks during emergency response [3, 12]. Simply put, social networks exhibit the ability to correct and filter out incorrect information through the natural collaborative process of discussion and consensus development.

![Figure 14](image.png)  
**Figure 14.** % of alerted students with incorrect info.

Due to the large number of group messages received by students, it appears to be beneficial to maintain a healthy skepticism of their content. The lower rates of group message belief seems to have suppressed the amount of incorrect belief generated by wrong messages or students’ misinterpretations of messages from people with whom they are not so familiar (figure 14).
6.3. Friend Messages (Scenario 2)
From the results, we can see that friend networks would need to include more close connections in order for systems based on direct personal messages to be effective at quickly and consistently spreading situational awareness. Less than half of students in Scenario 2 were alerted by messages (figure 5) and overall alert numbers were consistently lower (figure 3). This helps to confirm the intuition that social networks with high probability of incorrect or unclear information may not benefit from systems based solely on friend messages.

Friends’ info is trusted in the model more often than that of groups; this is apparently not always beneficial, as the level of both correct and incorrect information adoption were higher than for group messages. While they are more trusted, and less likely to be misinterpreted, they still maintained higher levels of incorrect info maintenance than other scenarios (figure 14). Also, the average proportion of students using group messages who had the correct information at the end of the attack was ~74%, while the proportion using friend messages was only ~67% (table 2). In the end, the lower message count and apparent danger in consistently placing high trust in friends seems to outweigh (or impair) the ability of the friend network to self-update as effectively as seen in other scenarios.

As the simulation progresses in Scenario 2, we can see that the rate of incorrect info update from friend messages stays lower while the rate of correct info update steadily rises (figures 11 and 13); the opposite phenomena can be seen for Scenario 4. While an initial conclusion might be that the inclusion of group messaging in Scenario 4 may be responsible for the difference, we know this is probably inaccurate as group messages in Scenarios 3 and 4 exhibit similar behavior to one another in this respect (figures 10 and 12). An alternative explanation lies in the fact that in the current model students will not update their belief to the same belief; if they receive a message that confirms what they already know, they simply ignore it. In this case, we can reason that the differences and somewhat mirrored behavior in Scenarios 2 and 4 are due to pre-existing beliefs not being updated and thus producing lower plot lines.

An influential factor here may also be that friend messages are considered first; conversely, a student’s existing belief in Scenario 4 may be due to the adoption of belief from a previous group message, which would have been considered last and may not have been altered. However, friend networks seem to be more susceptible to losing that correct understanding due to the increased (and apparently sometimes misplaced) trust in friend messages.

6.4. Further Comparison of Different Scenarios
Despite the inclusion of the more sporadically received (yet more trusted) friend messages, the results did not significantly differ between Scenarios 3 and 4 (figures 3, 4, 5, 6, 7, 10 and 12). Also, the results from Scenario 2 largely reflects the trends seen in Scenario 1; despite producing more alerted students, the still-slow rise in the number of aware students led to the same (slower) event reporting that we saw in Scenario 1. The end-of-run results are telling as well; it seems that while the inclusion of group messages has a significant impact on the outcome, the inclusion of friend messages has meaningfully smaller impact (table 2). Scenarios 3 and 4 produced at least 8% higher levels of correct information held by students compared to Scenarios 1 and 2, and led to about 10 fewer casualties than scenarios without group messages. These figures seem not to be significantly impacted by friend messaging ability.

In looking at the figure for the percent of alerted students with incorrect info (figure 14), we can also see that the value for Scenario 3 is often lower than that in other scenarios – often significantly lower than those in Scenarios 1 and 2. While the Scenario 3 values are often comparable to Scenario 4’s, in figure 14. Scenario 3 seems to consistently correct the initial confusion resulting from attacker movements earlier on in the simulation, and at a faster rate, than Scenario 4. We should note that this is usually the time period that the first event reports are being made in. In terms of alert numbers, all four scenarios show approximately the same results for around 1.5 minutes into the attack (figure 3). After that, the number of students alerted in Scenarios 3 and 4 quickly explodes, with a difference becoming...
quickly greater 500 students (which is ~25% of the on-campus student population) in comparison to Scenarios 1 and 2.

7. Conclusion and Suggestions for Future Work
In the current project, we have looked at the effects of messaging under an information sharing system structured similarly to the popular WeChat platform. Our results indicate that individuals’ information sharing and belief dynamics and overall information sharing system structure can have a significant impact on the proportion of students who are both aware of an attack and able to maintain correct information about the situation as it evolves.

While preliminary, our results indicate that emergency communication systems based on direct private messages between people may not be optimal for effective emergency response due to their inefficiency at updating people on the scale needed in the situation. Group messages may be more efficient at causing a meaningfully earlier event report time due to the larger numbers of alerted people they produce. While it seems that placing so much faith in our friends isn’t necessarily a good thing, a healthy level of skepticism in group chats (that is, an appropriate level considering all aspects of the situation) may actually be helpful in maintaining the ability of the group to effectively produce accurate situational understanding in its members.

More work in this area should be conducted. But we have seen here that information sharing systems and behaviors play a significant role in campus attack response, as they do in other areas of emergency response. As the majority of information sharing systems currently used in emergency response are not designed for this purpose, we would recommend that future research also be conducted to consider ways in which we can learn from the behavior and results seen in existing systems, in order to develop more effective purpose-driven emergency information sharing systems. Based on the current study, it seems that consideration of ways that the benefits of group chats can be utilized, while addressing factors that may limit their effectiveness, may be a promising research pathway.

7.1. Further Suggestions for Future Work
As the current work is an initial exploration into this area, naturally there will be ways to improve which might yield better results. One example is the probability of an inaccurate/incomplete message; further work could be done to fine-tune our assumptions, or study the effects of info sharing systems as the probability changes. Also, we currently consider each message independently of others; a better approach would be one that considers the content of multiple messages when making an information belief decision, and perhaps gives more weight to some messages based on how frequently their content has been repeated. There is also the issue of group chats; a more complex model of group chat membership should be considered in future work. An interesting project might be to consider the effect that the topology of social networks based on group chats might have on students’ attack response; as we have seen here in Scenario 2, more degrees of separation between network nodes translates to a less effective alert and update process.

One thing that we should note is that while the model accounts for the possibility of students fighting back against the attacker, we did not include this possibility in our tests as we wanted to focus on information sharing. In this project, we also did not code our simulated students to change their evacuation course according to their currently believed attack location. We believe that this may play a significant part in why we saw a steady linear increase in the casualty level, with similar slope (~1) across the different scenarios we tested (figure 2). We would expect that if this change is made, we will see more fluctuation in the casualty rate, and possibly in the overall number of casualties as well.

Follow-up work should also consider the effect of dynamics that might result from different information sharing system structures. For example, this model could be adjusted to look at how response might be improved by simply raising awareness of the emergency reporting phone number and using an official emergency notification system – or if a system intentionally designed for on-campus emergencies might provide better results.
Acknowledgment
This research was funded by the National Natural Science Foundation of China (Grant No. 71971127) and the Hylink Digital Solutions Co., Ltd. (120500002).

References
[1] Abedin B and Babar A 2018 Institutional vs. non-institutional use of social media during emergency response: A case of twitter in 2014 Australian bush fire Information Systems Frontiers 20 (4) 729-740.
[2] Hughes AL and Palen L 2009 Twitter adoption and use in mass convergence and emergency events International Journal of Emergency Management 6 (3-4) 248-260.
[3] Thomson R, Ito N, Suda H, Lin F, Liu Y, Hayasaka R, Isochi R and Wang Z 2012 Trusting tweets: The Fukushima disaster and information source credibility on Twitter Proceedings of the 9th International ISCRAM Conference pp 1-10.
[4] Linger K 2018 Analysis of the Police Response to Mass Shootings in the United States between 1966 and 2016 (University at Albany) Bachelor Thesis.
[5] Zhang F, Wang S and Song Z 2018 Evacuation during violent attacks: agent-based modeling and simulation 2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS) pp 412-415.
[6] Ma L, Chen B, Qiu S, Li Z and Qiu X 2017 Agent-based modeling of emergency evacuation in a railway station square under sarin terrorist attack International Journal of Modeling, Simulation, and Scientific Computing 8 (02) 1750022.
[7] Xi J Y S and Chan W K V 2019 Simulation of knife attack and gun attack on university campus using agent-based model and GIS Proceedings of the 2019 Winter Simulation Conference pp 263-272.
[8] Briggs T W and Kennedy W G 2016 Active shooter: An agent-based model of unarmed resistance Proceedings of the 2016 Winter Simulation Conference pp 3521-3531.
[9] Anklam C, Kirby A, Sharevski F and Dietz J E 2015 Mitigating active shooter impact: Analysis for policy options based on agent/computer-based modeling Journal of Emergency Management 13 (3) 201-216.
[10] Okaya M and Takahashi T 2011 Human relationship modeling in agent-based crowd evacuation simulation International Conference on Principles and Practice of Multi-Agent Systems pp 496-507.
[11] Lee J Y 2019 Agent-Based Modeling to Assess the Effectiveness of Run Hide Fight (Purdue University) Master Thesis.
[12] Hiltz S R, Gonzalez J J and Van de Walle B 2012 Assessing and improving the trustworthiness of social media for emergency management: A literature review Proceedings of the Norwegian Information Security Conference (NISK) pp 19-21.