myjournal manuscript No.
(will be inserted by the editor)

Empirical Evaluation of Computational Fear Contagion Models in Crowd Dispersions

Jason Tsai · Emma Bowring · Stacy Marsella · Milind Tambe

Received: date / Accepted: date

Abstract In social psychology, emotional contagion describes the widely observed phenomenon of one person's emotions being influenced by surrounding people's emotions. While the overall effect is agreed upon, the underlying mechanism of the spread of emotions has seen little quantification and application to computational agents despite extensive evidence of its impacts in everyday life.

In this paper, we examine computational models of emotional contagion by implementing two models ((Bosse et al, 2009b) and (Durupinar, 2010)) that draw from two separate lines of contagion research: thermodynamics-based and epidemiological-based. We first perform sensitivity tests on each model in an evacuation simulation, ESCAPES, showing both models to be reasonably robust to parameter variations with certain exceptions. We then compare their ability to reproduce a real crowd panic scene in simulation, showing that the thermodynamics-style model (Bosse et al, 2009b) produces superior results due to the ill-suited contagion mechanism at the core of epidemiological models. We also identify that a graduated effect of fear and proximity-based contagion effects are key to producing the superior results. We then reproduce the methodology on a second video, showing that the same results hold, implying generality of the conclusions reached in the first scene.

Keywords Emotional contagion · Emotion modeling · Simulation

Jason Tsai
University of Southern California, Los Angeles, CA 90089
E-mail: jasontts@usc.edu

Emma Bowring
University of the Pacific, Stockton, CA 95211
E-mail: ebowring@pacific.edu

Stacy Marsella
USC Institute for Creative Technologies, Playa Vista, CA 90094
E-mail: marsella@ict.usc.edu

Milind Tambe
University of Southern California, Los Angeles, CA 90089
E-mail: tambe@usc.edu

1 Introduction

Emotional contagion, the tendency for one's emotions to reflect the emotions of others, has been shown to arise in a wide range of scenarios in everyday life (Hatfield et al, 1994). Its effects are felt every time someone cheerfully walks into the room with a big smile and brightens up everyone's day. Extensive work has been done in researching emotional contagion's role in occupations that require an employee to promote certain emotions in clients via displayed emotions, such as bill collectors promoting anxiety or flight attendants creating good cheer (Grandey, 2000; Pugh, 2001). Less often, but with far more severe implications, it is also felt during the spread of fear and anxiety that surrounds any crowd-based disaster.

Virtual agents designed for these domains must also incorporate the effects of emotional contagion. For example, virtual patients in clinical training applications must incorporate not only the linguistic response of a real patient to a clinician's questions (Kenny et al, 2008) but also a real patient's emotional response to a clinician's demeanor that results from emotional contagion. Similarly, an evacuation training simulation must include not only emotional contagion between simulated agents and its impact on escape behavior, but also exhibit emotions accurately to the user to mimic the contagion effects in a true evacuation (Tsai et al, 2011). With the growing awareness of the powerful impact that emotion has on human behavior, the contagion of these emotions can no longer be marginalized in virtual agents and must be accurately modeled and incorporated.

Recent work has sought to quantify the qualitative findings of social psychology into useable models, primarily drawing from two bodies of research on similar phenomena. Researchers at VU University introduced one of these in (Bosse et al, 2009b) (ASCRIBE) that used a deterministic, interaction-based model derived directly from a social psychology theory of emotional contagion (Barsade and Gibson, 1998). This model is a prototypical example of the heat dissipation phenomena studied in thermodynamics wherein neighboring substances will transfer energy to each other at rates unique to each substance (i.e., specific heat). In contrast, Durupinar (Durupinar, 2010) used a probabilistic threshold model wherein successive interactions with emotionally 'infected' people raises the chance of infection with an emotion. This model is a standard one from the extensive epidemiology literature that models the spread of diseases (Dodds and Watts, 2005; Kermack and McKendrick, 1927; Murray, 2002), the research in diffusion of innovations (Rogers, 1962), and social contagion work (Schelling, 1973).

Although both models come from studies of contagion phenomena, they use fundamentally different mechanisms. While work could proceed using both approaches by extending existing models to accurately reproduce increasingly complex situations, it remains unclear which contagion paradigm should be used in emotional contagion. Perhaps a new mechanism should be designed, but the lack of data in this domain makes evaluation very difficult. We not only empirically compare these two paradigms but begin to identify the key features that should be added to the underlying contagion mechanisms to further improve their fidelity in reproducing human emotional contagion.

Feature	ASCRIBE	Durupinar
Emotion Level	Continuous	Binary
Fear Level Impacts Contagion	Yes	No
Emotional Decay	No	Yes
Interaction Type	Individual	Threshold
Interaction Determinism	Deterministic	Probabilistic
Proximity	Yes	No

Table 1 Key model differences

As we outline in Table 1, there are 6 primary differences between the Durupinar and ASCRIBE models. First, the Durupinar model follows in the tradition of epidemiology where it is nonsensical to discuss a ‘degree’ of infection and uses a binary specification for emotional level, whereas the ASCRIBE model uses a continuous description. Second, with no degree of emotion, the Durupinar model cannot specify differences in contagion that may result from differing levels of emotion (e.g., a high-fear may cause nearby agents to become more fearful than a low-fear agent would). Third, the Durupinar model includes a decay factor whereas the ASCRIBE model does not. Fourth, the ASCRIBE model uses an individual interaction model where each agent encountered causes some contagion, whereas the Durupinar model uses a threshold model in which each encounter causes an increased chance of contagion. Fifth, the ASCRIBE model uses a deterministic interaction scheme, whereas the Durupinar model uses a probabilistic one. Finally, the latest ASCRIBE model (Bosse et al, 2011) incorporates proximity’s effect on contagion, whereas the Durupinar model does not.

We begin by using the ESCAPES evacuation simulation (Tsai et al, 2011) to explore the impact of replacing the original ESCAPES model with these two models on predicted outcomes, showing substantial differences in their predictions, motivating the need for an accurate model of emotional contagion in this context. Even in simulation, we are able to identify key differences that indicate epidemiological / social contagion models are less suited to modeling emotional contagion. Next, we attempt to reproduce a subset of 35 people from real video footage of a panic situation using each of the models, showing the ASCRIBE model to indeed be superior to both the Durupinar model and the original ESCAPES model, beating out the Durupinar model by 14% *per agent per frame* during the 15s scene. To identify which of the key features causes the differences in the results, we test hybrid models to conclude that while adding a ‘decay’ feature (as found in the Durupinar model) to the ASCRIBE model does not improve it, removing proximity effects and fear’s graduated effect on speed substantially worsen the model. Finally, we perform the same evaluation on a second video, extracting 10 people, and show the ASCRIBE model to again be superior, outperforming the Durupinar model by 12% per agent per frame during the *four-second* scene.

2 Related Work

Seminal works in social psychology first began the discussion around emotional contagion. In particular, (Hatfield et al, 1994) first codified the observed phenomena that were just beginning to receive researcher attention. Follow-up work by the co-authors as well as in related fields such as (Barsade and Gibson, 1998; Grandey, 2000; Pugh, 2001) in managerial and occupational sciences continued to detail the effects of the phenomenon in new domains. Recently, there have been works beginning to quantify emotional contagion and explore cross-cultural variations in attributes that affect emotional contagion (Doherty, 1997; Lundqvist, 2008).

From a computational perspective, the previously mentioned work from VU University and Durupinar are two of the most recent models of emotional contagion upon which a few follow-up works have been based (Bosse et al, 2009b,a, 2011). As mentioned in Section 1, the ASCRIBE model resembles heat dissipation models found in basic physics wherein each substance has its own heat dissipation rate and heat absorption rate. The Durupinar model draws inspiration from a long line of contagion models (Dodds and Watts, 2005; Kermack and McKendrick, 1927) that was popularized in the diffusion of innovations (Rogers, 1962) literature and has also seen heavy use in other types of social (e.g., belief, behavior, idea) contagion (Schelling, 1973).

Diffusion modeling in computer science has primarily followed the influence maximization paradigm formalized by (Kempe et al, 2003), where an entity seeks to maximize the diffusion of its influence across a known network. Follow-up works have considered numerous variations on both the linear threshold model (as in Durupinar) and independent cascade models of diffusion as well as competitive diffusion modeling where multiple parties compete to spread their influence across a network (Leskovec et al, 2007; Borodin et al, 2010; Chen et al, 2010). The primary focus of this type of work has been the identification of the optimal set of ‘seed’ nodes from which to begin a diffusion process to maximize the coverage of the spread. Typical example domains cited include social marketing and rumor spreading, but these works have not considered the diffusion of emotions in a crowd, nor used live crowd data for model verification or calibration as we do here.

3 ESCAPES

Although not the focal point, the ESCAPES evacuation simulation (Tsai et al, 2011) serves as the test bed for our models of emotional contagion, so we describe it briefly here. ESCAPES focuses on the features identified by experts that particularly affect airport evacuations (Diamond et al, 2010). It models realistic agent knowledge about the environment by only letting agents recall the location of the exit they entered from, which will not always be the nearest exit. Furthermore, agents have a random probability to forget the location of known exits, as has been noted to occur in the evacuation literature (Chertkoff and Kushigian, 1999). ESCAPES also models a realistic spreading of knowledge about the need for agents to evacuate, which is unlike most evacuation simulations in which all agents immediately know they must make their way out. This mimics the pre-evacuation delay that is widely studied in evacuation research (Mileti and Sorensen, 1990; J.L.Bryan, 2002). In airport settings, the presence and unique behavior of families was noted to be of particular importance (Diamond et al, 2010), so ESCAPES also models family groups that, if split up, will seek each other out before attempting to find an exit. Clearly, the presence and behavior of authority figures is also a key feature in airports and ESCAPES includes authority agents that exhibit a variety of patrolling behaviors, know the location of all exits in the environment, are notified earliest about the need to evacuate, and communicate their knowledge to surrounding agents.

Finally, the ESCAPES model also includes a baseline model of individual emotion and emotional contagion between agents. Specifically, we model individual fear and its impact on behavior by increasing the speed of more fearful agents. ESCAPES uses a basic model of emotional contagion that we use as a baseline for comparison with the ASCRIBE and Durupinar models. In the ESCAPES model, agents inherit the highest fear level of neighboring agents and maintain their fear level until they escape the environment. A full listing of the features, both first and second order, included and references to work citing their importance during evacuations can be found in Table 2. We replace the original model of emotional contagion with the Durupinar and ASCRIBE models to compare the simulated effects of using each of the different emotional contagion models. Specifically, we will show results examining the spread of emotion through the crowd and predicted evacuation safety metrics.

Phenomenon	Reference
People forget their entrance	(Chertkoff and Kushigian, 1999)
First-time Visitors	(Diamond et al, 2010)
Heightened emotions lead to chaos	(Smith and Ellsworth, 1985)
Herding behavior	(Helbing et al, 2000)
Pre-evacuation delay	(Mileti and Sorensen, 1990; J.L.Bryan, 2002)
Families gather before exiting	(Proulx and Fahy, 2008)
Authorities calm people	(Smith and Ellsworth, 1985)

Table 2 Phenomena modeled in ESCAPES

4 ASCRIBE model

Introduced in 2009 by researchers at VU University (Bosse et al, 2009b) and built upon in multiple works including (Bosse et al, 2009a, 2011), the ASCRIBE model iterates through all agents and deterministically calculates new emotional levels based on a set of individual and pairwise parameters that we describe here. The mechanism used resembles heat dissipation modeling in physics, wherein each material has a specific heat capacity, which can be likened to a person's susceptibility to other people's emotions in emotional contagion. As such, the model moves a crowd towards a weighted-average of the group's emotional levels, just as heat will dissipate until adjacent temperatures are the same, barring generative heat sources.

The model defines 5 parameters (shown in Table 3) for every pairwise interaction based on theory put forth in (Barsade and Gibson, 1998): level of sender's emotion

level of the sender's emotion	q_S
level of the receiver's emotion	q_R
sender's emotion expression	ϵ_S
openness for received emotion	δ_R
strength of the channel from sender to receiver	α_{SR}

Table 3 Aspects related to a sender *S*, receiver *R*, or both

q_S, level of receiver's emotion q_R, sender's expressiveness ϵ_S, receiver's openness δ_R, and the channel strength between S and R α_{SR}. All values are numbers in the interval $[0, 1]$. At each time step, each agent calculates the average emotional transfer from all relevant agents. Specifically, the differential equations for emotional contagion in a group G of agents is:

$$dq_R/dt = \gamma_R(q_R^* - q_R)$$

for all $R \in G$, where γ_R is the overall strength at which emotions from all other group members are received, defined by $\gamma_R = \Sigma_{S\in G\setminus\{R\}}\gamma_{SR}$. q_R^* is the weighted combination of emotions from the other agents, defined with a weight factor:

$$w_{SR} = \epsilon_S\alpha_{SR}/\Sigma_{C\in G\setminus\{R\}}\epsilon_C\alpha_{CR}$$

$$q_R^* = \Sigma_{S\in G\setminus\{R\}}w_{SR}q_S$$

Specifically, from a sender S to a receiver R, the strength of the emotion q_S received would be $\gamma_{SR} = \epsilon_S \cdot \alpha_{SR} \cdot \delta_R$. (Bosse et al, 2009b) details the mathematical formulation, but the emotional level of an agent converges towards a weighted average of the group's emotional level. The speed at which this convergence occurs as well as the weighting depend on the parameter settings for the channel strength, expressiveness, and openness for each agent as well as, of course, their individual emotional levels.

The latest version of the model (Bosse et al, 2011), extends the original emotional contagion model and includes beliefs and intentions and belief/intention contagion as well. However, as our goal is to empirically evaluate emotional contagion models and the latest work extends far beyond simply emotional contagion, we leave its validation to future work. Thus, we do not use the extended model but instead modify the initial model by incorporating a proximity effect as done in (Bosse et al, 2011).

5 Durupinar Model

Durupinar (Durupinar, 2010) uses a probabilistic threshold model based on epidemiological models of disease contagion. While many types of epidemiological models exist (Dodds and Watts, 2005; Kermack and McKendrick, 1927; Murray, 2002; Schelling, 1973), Durupinar opts for a baseline model from (Dodds and Watts, 2005). In this model, individuals can be in either *susceptible* or *infected* states. Other models incorporate additional states such as inoculated, recovered, etc. which could be incorporated in extensions to Durupinar's basic model, but have not been explored in the context of emotional contagion. The epidemiological model's applicability to emotional contagion was not discussed in (Dodds and Watts, 2005), from which Durupinar drew, but its use by Durupinar assumes similarity between disease spread and emotion spread that we criticize in this work.

Each agent begins with a randomized threshold drawn from a pre-determined log-normal distribution. At each time step, T, for each agent, a random agent is chosen from the relevant population group. If the agent is infected, it generates a random dose, d_j, with size drawn from a pre-determined log-normal distribution and passes it to the original agent. If the agent is not infected, then a dose of 0.0 is generated. Each agent maintains a running history of the last K doses received. If the cumulative total of all doses in the agent's history exceeds his threshold, the agent enters the infected state (Equation (1)). This causes the emotion level to be set to 1.0 with an exponential decay towards 0.0 at a rate β (Equation (4)), at which point the agent re-enters the susceptible state. A non-zero emotion level indicates that the agent has the emotion, but the actual value does not hold meaning other than to track the decay. The random dose and threshold are generated from log-normal distributions (Equations (2),(3)) with user-specified averages and standard deviations and K is a static global variable.

$$D_j(t) = \sum_{t'=t-K+1}^{t} d_i(t') \tag{1}$$

$$d_j = \text{log-}\mathcal{N}(\mu_{d_j}, \sigma^2_{d_j}) \tag{2}$$

$$T_j = \text{log-}\mathcal{N}(\mu_{T_j}, \sigma^2_{T_j}) \tag{3}$$

$$e_t = e_{t-1} - \beta \cdot e_{t-1} \tag{4}$$

Durupinar also provides a psychological basis for setting the dose and threshold distribution values by incorporating findings from Jolliffe and Farrington (Jolliffe and Farrington, 2006) on the correlation between the basic empathy scale and the OCEAN personality factors. A much richer emotional model is also described, but for the purposes of this study, we only use the emotional contagion model. The particular model introduced here is but one example from the range of similar contagion models (Kermack and McKendrick, 1927; Murray, 2002; Schelling, 1973), but they all share a binary, probabilistic treatment of effect. While it may seem trivial to interpret the decaying emotional indicator as a continuous variable, this alteration proves unhelpful in our experiments. As we show in the following sections, this fundamental difference between the heat dissipation-style models (such as the ASCRIBE model) and epidemiological models leads to inaccuracies in the Durupinar's modeling of emotional contagion.

6 Simulation Experiments

Although they are similar from a computational performance perspective, the ASCRIBE model and the Durupinar model use very different mechanisms to recreate emotional contagion. Thus, we evaluate the impact of these differences in two ways,

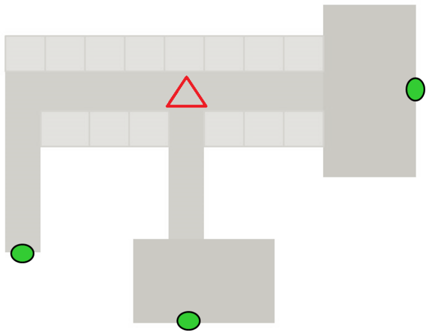


Fig. 1 Evacuation scenario.

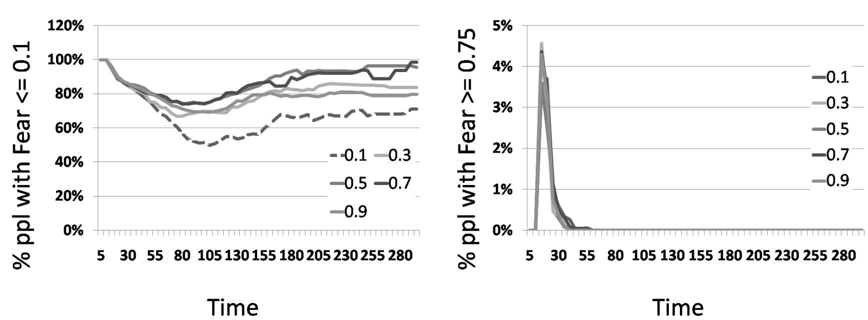
beginning first with simulation. We ran the evacuation simulation, ESCAPES, using each model to perform sensitivity analysis as well as identify any qualitative trends that might support or discredit either one of the models. We can also evaluate the model's robustness to errors in parameter estimation, which is extremely important in emotional and crowd modeling which usually lack high fidelity, fine-grain data.

For all the experiments discussed in this section, the same map was used (spatial layout can be seen in Figure 1) and 30 trials were run for each setting. It features 2 large spaces that represent airport boarding areas, each with an exit (marked with dots), connected by hallways which are lined with smaller spaces that represent shops. 15 seconds into the simulation, an event occurs at the center of the scenario (marked by the triangle), inciting fear and a need to evacuate that is communicated by authority figures to pedestrians. For initial fear levels, we define a 'seeing distance', σ_d. Agents within this distance of an event will immediately have a fear level of 0.75 in the ASCRIBE model and 1.0 in the Durupinar model, since the Durupinar model does not feature a continuous measure of fear. We also define a 'hearing distance', ω_d, within which the agent will receive 0.1 in the ASCRIBE model and 1.0 in the Durupinar model. The scenario features 100 normal pedestrians, including 10 families of 4 each, as well as 10 authority figures that patrol the scenario. In Sections 6.1 and 6.2 we evaluate model robustness and then identify qualitative differences in Section 6.3.

6.1 ASCRIBE model

In examining the contagion effect, the parameters of interest in the ASCRIBE model were the individual expressiveness settings and individual openness settings. The channel strength is set to 1 if an agent is nearby and 0 otherwise, as done in (Bosse et al, 2011). Given that we had a whole population of agents, we elected to use randomly drawn values for expressiveness and openness based on a normal distribution. It may be useful to relax this assumption of a normal distribution in future work, perhaps when more quantitative analysis of human contagion parameters becomes available. We explored variations of the averages and standard deviations (SD) used, but surprisingly, none yielded qualitative changes in the trends observed in the simulation from both a contagion perspective (i.e., how the fear spread) and a safety analysis. The only exceptions were, unsurprisingly, when the receiver openness or sender expressiveness parameters varied tightly around a very low mean, often leaving many agents with 0.0 openness or expressiveness. Very low receiver openness values created perpetual high-fear sources that constantly raised the fear levels of surrounding agents without ever dissipating their own fear. At very low sender expressiveness values, the majority of agents remain at their initial fear level. Both cases result in vastly different trends from the mean convergence behavior seen in the other settings.

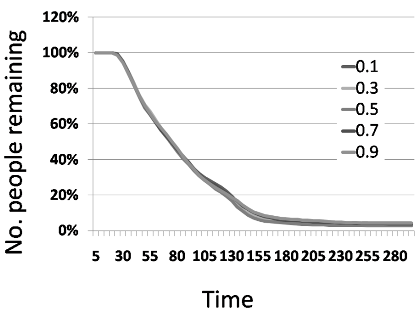
To illustrate the contagion effect of variations in the parameter settings, Figure 2a plots the percentage of people with low fear (≤ 0.1) on the y-axis and the time step on the x-axis, while Figure 2b shows the same results for high-fear people. In both figures, openness varied from 0.1 to 0.9 in increments of 0.2 while keeping a SD of 0.1 and sender expressiveness was fixed with an average of 0.5 with a SD of 0.1. In Figure 2a, when an event first occurs, those near it become fearful and slowly raise nearby peoples' fear as they move towards exits, causing a steady decline in the percentage of people with fear less than 0.1 that only rises again as fearful agents make their way out of the simulation. Note how the dotted line (0.1) dips much lower than the other lines, showing the exception mentioned above. In Figure 2b, a few agents near the event have their fear raised very high, but as they encounter zero-fear agents, their fear levels are brought down below 0.75 and never again rise higher since no new events occur. The tightness of the lines implies that the trend is robust to variations in the average receiver openness except at very low settings. Similar tightness of lines was observed in variations of sender expressiveness, with the same exception.



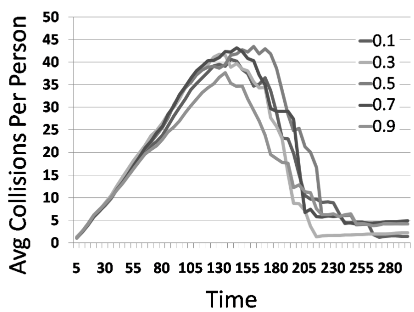
(a) Percent low-fear agents (b) Percent high-fear agents

Fig. 2 ASCRIBE model: Variations of Openness on contagion

Figure 3 focuses only on variations of openness to illustrate the trends observed in evacuation safety. Figure 3a plots the percentage of people that have not escaped on the y-axis and time steps on the x-axis, whereas Figure 3b shows the average number of agent-agent collisions accumulated by each person remaining in the simulation on the y-axis and time steps again on the x-axis. Since the simulator does not explicitly handle agents colliding, we define 'collisions' as anytime two agents touch each other. Figure 3a shows all parameter settings for openness leading to almost identical escape times for people during the simulation. Figure 3b shows extremely similar collision counts for people across the parameter space as well. Thus, when measuring the second-order metrics of pedestrian escape times and number of collisions, the model remains robust to parameter variations of the type tested. Variations of the other parameters' averages and standard deviations all resulted in the same extremely tightly clustered lines as seen in Figures 2 and 3 (with the previously noted exception).



(a) Percent pedestrians unevacuated



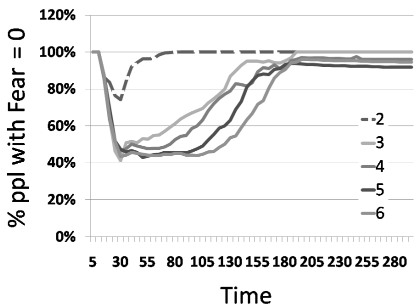
(b) Average collisions per person

Fig. 3 ASCRIBE model: Variations of openness on safety

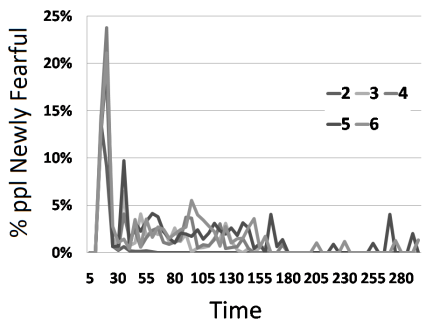
6.2 Durupinar Model

Sensitivity analysis of the Durupinar model is considerably more delicate than the ASCRIBE model, because although there are only 5 key parameters for the whole population (as compared to 2 per individual plus 1 for each pair for the ASCRIBE model) and even a small change in a single parameter can cause extreme changes in the trends observed. Thus, we begin with experimentally chosen default values and vary each parameter to identify key sensitivities. In particular, we begin with a baseline of K of 4, dose average of 2, dose standard deviation of 0.5, threshold average of 7, and threshold standard deviation of 2.

Figure 4a shows the percentage of no-fear pedestrians ($= 0$) on the y-axis and time steps on the x-axis, with each line representing a different setting of K. Figure 4b shows the percentage of newly fearful pedestrians (defined as ≥ 0.75) during the same variations of K. Unsurprisingly, altering any one of the parameters' averages *or* standard deviations individually alters the magnitude of the contagion effect, but not the overall trends. The exceptions are at values far from the baseline. For example, at extremely low values for K or dose distribution average and at extremely high values for threshold distribution average, when very few agents become fearful at all, as seen in the dotted $K = 2$ line in Figure 4a. This implies that the model remains robust to parameter changes with respect to the contagion trends that emerge as long as parameter values are chosen within a tolerance of the baseline. Similar results were found for variations of threshold and dose strength averages and standard deviations.



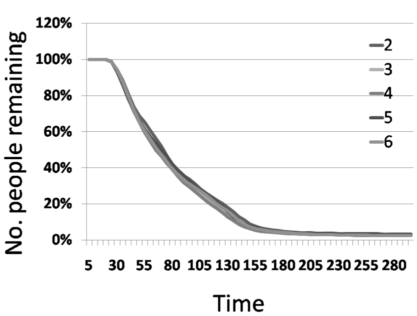
(a) Percent no-fear pedestrians



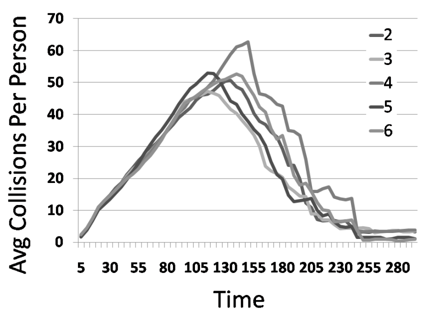
(b) Percent new-fear pedestrians

Fig. 4 Durupinar model: Variations of K on contagion

We again explored the second-order impacts of parameter variations on the safety of the evacuation by measuring the evacuation rates and average number of collisions of pedestrians in the simulation. Figure 5a shows the percentage of people that have not yet evacuated on the y-axis and the time steps on the x-axis, while Figure 5b shows the average number of agent-agent collisions accumulated for each person remaining in the simulation on the y-axis and the time steps on the x-axis. As in the ASCRIBE results shown previously, both of these graphs show extremely similar results across the parameter space tested. Variations of other parameters showed very similar results.



(a) Percent pedestrians unevacuated



(b) Average collisions per remaining person

Fig. 5 Durupinar model: Variations of K on safety

6.3 Key Differences

In Sections 6.1 and 6.2 we have shown the ASCRIBE model to be robust to parameter variations (except at the extreme of zero) and the Durupinar model to be robust if we stay within a tolerance of a baseline. In conducting these simulation tests and taking a closer look at the contagion effect, we already find that a number of key differences can be identified between the two models. One difference can be seen by comparing Figures 2b and 4b, where the spikes occurring throughout the graph indicate that Durupinar model produces newly fearful agents throughout the life of the simulation, *regardless of the nature of the event*, and the ASCRIBE model only exhibits a spike due to the impact of the event. Under the Durupinar model, fear can be transferred indefinitely under certain parameter settings. In the ASCRIBE model, encounters with agents who are less fearful will slowly erode the average fear level, eventually reaching zero after sufficiently many agents have been encountered.

Also, combining the binary fear metric with a speed modifier, as done in ESCAPES, results in only extremes of movement speeds. While one could argue that this is a result of the simulation, the model itself cannot incorporate any gradation of effect. For example, even if we directly map the fear level (as it decays) to the speed modifier, an agent that is near zero-fear (and is hence traveling slowly) can infect another agent who will then dart off at maximal speed since he begins at maximal fear, as evidenced by the spikes in Figure 4b. This may occur as a result of physiological or informational changes, but no evidence suggests this would occur from emotional contagion alone. A more fundamental alteration is needed to change this aspect of epidemiological / social contagion models for convincing application to emotional contagion.

Finally, as mentioned, the Durupinar model does not include a proximity of effect, whereas the ASCRIBE model does. This obviously means that the Durupinar model could potentially cause contagion of emotions to agents randomly throughout the world of the simulation, a very unrealistic effect, as emotional contagion requires *some* form of interaction by definition. As seen in a comparison between Figures 2a and 4a, the Durupinar model induces more fearful agents far more rapidly than the ASCRIBE model does because its contagion calculation incorporates the entire population immediately.

7 Scene Reproduction

Now we discuss the validation method used to evaluate the models of emotional contagion, first used in (Bosse et al, 2011). In their work, VU University researchers used a 15-second portion of a crowd panic scene in Amsterdam caused by a screaming person[1] as their dataset for validating their general mental state contagion model. In processing the data, the researchers traced the locations of 35 people scattered through the crowd through the 15 seconds, converted these into top-down coordinates and built a simulator to reproduce the paths of the people in simulation (for more detail on the spatial parametrization, please refer to Bosse et al (2011)). The 35

[1] http://youtu.be/0cEQp8OQj2Y

people chosen can be interpreted as point estimations of the speed and trajectory of subsets of the crowd throughout the scene. Thus, accurately predicting the movement of these individuals would translate into accurate prediction of the overall movement of the crowd. Furthermore, since realistic agent collisions remains an open research question, using every individual in the crowd would present novel challenges beyond the scope of this work.

The operating hypothesis was that a simulator without their mental state contagion model would not be able to reproduce the scene as accurately as a simulator with it. To test this hypothesis, the researchers tuned parameters associated with each agent's maximum speed, a global parameter specifying a 'sight range' within which agents could 'see' the event, and an initial desire to remain in place. The tuning was done via hill-climbing to minimize the error produced by the simulator, testing each parameter and moving a single parameter at a time in the direction of highest error reduction until a local optimum was reached. Error was defined as the sum of the average distances from each simulated agent to the corresponding real people's locations over the life of the simulation. Then, they incorporated the mental state contagion model, tuning a parameter associated with the proximity of contagion and showed that lower error was achieved with this addition.

We extended the approach of the VU University researchers by importing the 35 agent traces into the ESCAPES simulator and setting 3 exit locations towards which agents proceed when the simulation starts. The locations were chosen to roughly mimic the real situation, leading to most agents moving in the same direction as the people did. Some agents did not move precisely in the simulated direction as a result of obstructions that we did not model and a person very close to the screaming person that barely moved. The primary task was to match the crowd's location over time, first without contagion effects and then with each contagion model in turn. Since people's directions did not vary based on the emotion, the contagion model could only impact the speed of each agent.

The speed of an agent, without incorporating contagion effects, is based on the emotional level multiplied by the maximum speed multiplied by a distance-based modifier. The distance-based modifier is σ_s if the agent is within sight range and ω_s if the agent is only within hearing distance. We include these tunable speed modifiers so that the simulation is robust to the choice of initial fear levels, which is particularly helpful given the lack of data surrounding how to set the initial fear levels.

As an example of the speed calculation, under the ASCRIBE model, if an event occurred within hearing distance but not seeing distance of an agent, and the hear-range speed modifier was 0.2, the agent's speed would be $(0.1)(0.2)(S_{max})$, where S_{max} is the maximum speed allowed in the simulation. Under the Durupinar model, the 0.1 would be replaced with a 1.0. σ_d, ω_d, σ_s, and ω_s are global parameters applied to all agents that we tune experimentally via the same methodology as used in (Bosse et al, 2011), where our measure of error is the sum of simulation-space coordinate distance between the simulated agents and the actual agents.

For each contagion model, we use the default settings discussed in Section 6, with the exception of the ASCRIBE model's channel strength, which we set to 1.0 or 0.0 depending on the proximity of other agents, as was done in (Bosse et al, 2011). In the ASCRIBE model, we follow (Bosse et al, 2011) and fix Receiver Openness and

Sender Expressiveness each to 0.5 *for every agent*, but allow the proximity parameter to be tuned. In the Durupinar model, we set the dose history to 6, the mean and standard deviation of the dose strength distribution to 2 and 0.5, and the mean and standard deviation of the threshold distribution to 7 and 2. The ESCAPES contagion model, used as a baseline for comparison, only requires tuning of the proximity parameter as it simply brings all agents to the highest level of fear found in surrounding agents. In an attempt to not only identify which model is more appropriate but also to discern key features from unsupported augmentations, we used each model as given, then turned on/off implementations of 'decay', emotional level impacting speed, and proximity effects. For each parameter setting, 30 trials were run.

7.1 Amsterdam Crowd

We first use the Amsterdam crowd scene featured in (Bosse et al, 2011). In their results, VU University researchers found that the inclusion of contagion effects achieved significantly less error in reproducing the movement of a selection of 35 agents from the crowd scene. Upon closer inspection, the data revealed that the subset of agents within a particular radius surrounding the event caused the majority of every model's error. Specifically, approximately 80% of the error in each of the models' results can be attributed to the 13 agents nearest the explosion. The distinction between 'near' and 'far' agents is an empirical categorization based on the error attribution mentioned.

We show the error breakdown in Figure 6. Three categories of error are shown: faraway agents, the agent closest to the yelling, and the other nearest agents excluding the closest agent. The agent closest to the yelling barely moved in the video, which is a situation that the cognitive model of ESCAPES does not naturally simulate. Hence, all models produce large errors quite unrelated to the underlying emotional contagion model. The faraway agents, by contrast, move extremely little, making it easy to fit any model to them by simply forcing those agents to remain completely still. Thus, the largest portion of the error, that caused by the agents near the event (except the closest agent) also provides the most potential for the emotional contagion models to differ.

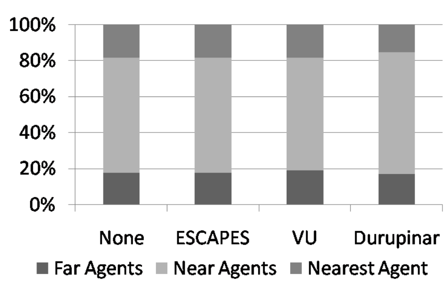


Fig. 6 Amsterdam crowd (35 agents): Error attribution

(a) Base models

Model	Overall	Near
None	0.375	0.699
ESCAPES	0.375	0.698
ASCRIBE	0.362	0.663
Durupinar	0.383	0.758

(b) ESCAPES variations

Model	Overall	Near
Base	0.375	0.698
Decay	0.379	0.703
No Speed	0.381	0.721
No Prox	0.385	0.721

(c) ASCRIBE variations

Variation	Overall	Near
Base	0.362	0.663
Decay	0.363	0.687
No Speed	0.387	0.767
No Prox	0.414	0.797

(d) Durupinar variations

Model	Overall	Near
Base	0.383	0.758
No Decay	0.387	0.771
Speed	0.388	0.784
Prox	0.380	0.754

Table 4 Amsterdam crowd: Average error (in pixels) per agent per frame

The results from the different variations of each model is listed in Table 4. Table 4a shows the results for the base models as defined previously, illustrating OVERALL error (for all 35 agents) as well as the error associated with the most substantial group of agents, the 12 NEAR the event, excepting the closest agent. Table 4b shows the variations associated with the original ESCAPES formulation. The second line of the table indicates that a 'decay' feature was added to the base model. The third line indicates that we turned on/off the effect that different levels of fear have on speed. When off, this means that any level of fear causes agents to travel at maximum speed. When on, the speed of travel is proportional to the fear level. Finally, the fourth row represents whether the contagion effect was moderated with a tuned proximity effect. Tables 4c and 4d show the analogous set of variations for the ASCRIBE and Durupinar models.

No results from the ESCAPES contagion formulation were statistically significantly better than the No Contagion case. This, as well as all remaining statistical tests in this work, was measured with a one-tailed t-test. This indicates that the ESCAPES contagion model does not add anything in the context of this dataset. In sharp contrast, all results for ASCRIBE and Durupinar were statistically significantly different from the No Contagion case, although in the case of Durupinar, they were significantly *worse* ($p < 0.001$). As found in (Bosse et al, 2011), the ASCRIBE model's formulation provided substantial improvements in the simulation's ability to reproduce this scene (14% superior to Durupinar for NEAR agents in the Base cases for the 15s clip).

For ESCAPES, no feature change offered statistically significantly different results from the base case, implying that in this formulation, for this data set, adding 'decay' did not help and the presence of 'speed' and 'proximity' features did not add value to the model either. In the ASCRIBE model, adding 'decay', removing 'speed', and removing 'proximity' all had statistically significantly negative impacts on the results ($p < 0.001$). This implies that the 'speed' and 'proximity' features were crucial to generating the positive result in the Base case and adding 'decay' does not improve it. Finally, removing 'decay' produced significantly worse results in the Durupinar model, and the other two variations did not produce statistically different results.

These results imply that the ASCRIBE model's contagion mechanism and current formulation provides the highest fidelity in modeling this dataset versus other variations and models tested. To properly frame the magnitude of improvement, consider a crowd being modeled for five minutes. In real terms, the 14% average difference between ASCRIBE and Durupinar amounts to over *two meters* of error over the 12 NEAR agents in a *single frame*. 'Small' errors like this in the first 15s can easily snowball into a completely different crowd structure after five minutes, suggesting much larger implications to this 14% improvement.

(a) Amsterdam video (b) Greece video

Fig. 7 Amsterdam and Greece video screenshots

7.2 Greece Crowd

Since one dataset could be particularly well-suited to the ASCRIBE model, we elected to perform the same process on a second video from protests in Greece in 2010[2], where officers fired tear gas into the middle of a small crowd. The clip used was from 0:16 to 0:20, from which 24 frames were extracted for analysis. 10 figures throughout the crowd were traced for the duration of the clip. Conversion of the pixel coordinates into top-down coordinates was done by first estimating true axes in the top-down view by tracing the sidewalk and steps that were perpendicular to the sidewalk. Then, the distance to each of the axes was calculated (where 'distance' is measured from the point to the axis, parallel to the other axis) and used as the new coordinates.

Even in such a short video clip with such a small crowd that we are able to match extremely well, the emotional contagion models still showed significant differences. Surprisingly, the original ESCAPES model performs extremely well, matching the ASCRIBE model's accuracy. However, as before, we see the Durupinar model again performing substantially worse than all other models, implying some generality of

[2] `http://www.youtube.com/watch?v=NsoDwM_KKfo`, posted May 5, 2010

(a) Base models

Model	Error
None	1.635
ESCAPES	1.478
ASCRIBE	1.478
Durupinar	1.656

(b) ESCAPES

Model	Error
Base	1.478
Decay	1.474
No Speed	1.567
No Prox	1.658

(c) ASCRIBE

Variation	Error
Base	1.478
Decay	1.466
No Speed	1.653
No Prox	1.660

(d) Durupinar

Model	Error
Base	1.656
No Decay	1.653
Speed	1.669
Prox	1.654

Table 5 Greece crowd (10 agents): Average error (in pixels) per agent during the simulation

the previous result. In fact, this scene is an even stronger testament than the previous one, as the ASCRIBE model performs 12% better than Durupinar in the Base case per agent per frame during only a *four-second* clip as opposed to the 15s Amsterdam clip. For both the original ESCAPES model and the ASCRIBE model, removing fear's impact on speed and the proximity effect statistically significantly worsen's the model's accuracy ($p < 0.001$). Surprisingly, the ASCRIBE model benefits from the addition of a decay component ($p < 0.001$), implying that a decay effect may be context-dependent.

8 Conclusions

In this work, we have made the first attempt to compare existing models of emotional contagion and identify key attributes of appropriate models using real data. The ASCRIBE model produced a 14% improvement *per agent per frame* over the Durupinar model in a 15s clip and a 12% improvement in only a *four-second* clip. After attempts to transform the Durupinar model into one more similar to the ASCRIBE model with little success. These results were consistent across 45 total agents in two unrelated crowd dispersion scenes. This suggests that the primary cause of the statistically significantly worse performance found with the epidemiological / social contagion model is in the mechanism of contagion itself, which is probabilistic and uses a binary representation of the effect. Although the ASCRIBE model requires setting $(N^2 + N)$ parameters to model N agents, even when we do away with them by fixing openness/receptiveness and only formulaically varying channel strength, the model produces superior results, implying that the underlying heat dissipation-style mechanism is better-suited to the phenomenon. In actual crowd modeling, simulators could use population averages for the parameters, as found in recent work (Doherty, 1997; Lundqvist, 2008) (instead of arbitrarily setting them at 0.5), resulting in a simplified model with one or two global parameters similar to 'specific heat capacities' for people's emotional transfer strength and one formulaic descriptor of proximity's impact. This leaves a simple, data-driven model of emotional contagion with empirical evidence supporting its superior performance.

A desirable next step would be to conduct this analysis on more videos, however, data collection remains a critical open challenge. Controlled experiments with crowds, in which all data and parameters can be controlled, are typically not practical or ethical for phenomena of interest such as evacuations. Videos of uncontrolled crowds, on the other hand, rarely ever use a still camera with an aerial view. Even with clean videos, extracting the necessary data must be done by hand due to the noise in the data. These issues make it extremely difficult to calibrate model parameters with real data. However, potential solutions do exist and represent key avenues for future research. One is to immerse people in virtual environments and use virtual characters to substitute for the other people in a crowd. Another is to rely on the fact that as our streets become more commonly instrumented with a range of cameras for security and with people commonly carrying camera devices then fortuitous capturing of usable videos may become more common. Another possibility is to extrapolate behaviors from smaller interactions such as dyads and triads in an experimentally verifiable way. Developing such data collection and model calibration techniques is a key open problem in simulation research.

This work serves as a first step in addressing this challenge by presenting a technique for setting model parameters for emotional contagion models based on real data. By comparing against real data, we are able to hone in on the key model attributes that influence the speed and strength of emotional contagion. Armed with a deeper understanding of emotional contagion models, the design of virtual agents can more accurately mimic human responses to emotional situations in their interactions with other agents as well as humans. For example, virtual patients that understand questions and respond properly (Kenny et al, 2008) will also react to the user's smiles, nods and other facial/vocal features to train clinicians to control the emotional contagion they inevitably cause. Virtual agents in emergency response simulations will not only be able to exhibit appropriate behaviors for a trainee to view and interact with, but also have a more accurate emotional effect on the user that will prepare him/her for the psychological strains that will inevitably arise. Only with the comprehensive quantitative understanding of emotional contagion that we have begun developing here will we be able to produce truly interactive, human-like agents.

Acknowledgements We would like to acknowledge the generosity of the VU University group that we have cited extensively in this work for sharing their experience and resources pertaining to the video evaluation. This research was sponsored in part by the U.S. Army Research, Development, and Engineering Command (RDECOM) Simulation Training and Technology Center (STTC). The content or information presented does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred. This research was also supported by the United States Department of Homeland Security through the National Center for Risk and Economic Analysis of Terrorism Events (CREATE) under award number 2008-ST-104-000013. However, any opinions, findings, and conclusions or recommendations in this document are those of the authors and do not necessarily reflect views of the United States Department of Homeland Security, or the University of Southern California, or CREATE.

References

Barsade SG, Gibson DE (1998) Group Emotion: A View from Top and Bottom. In: Gruenfeld D, Mannix E, , Neale M (eds) Research on Managing on Groups and Teams, JAI Press, pp 81–102

Borodin A, Filmus Y, Oren J (2010) Threshold models for competitive influence in social networks. In: WINE, pp 539–550

Bosse T, Duell R, Memon ZA, Treur J, Wal CNVD (2009a) A Multi-Agent Model for Emotion Contagion Spirals Integrated within a Supporting Ambient Agent Model. In: PRIMA-09

Bosse T, Duell R, Memon ZA, Treur J, Wal CNVD (2009b) A Multi-Agent Model for Mutual Absorption of Emotions. In: ECMS-09, European Council on Modeling and Simulation, pp 212–218

Bosse T, Hoogendoorn M, Klein M, Treur J, van der Wal N (2011) Agent-Based Analysis of Patterns in Crowd Behaviour Involving Contagion of Mental States. In: IEA/AIE-11, Springer Verlag

Chen W, Wang C, Wang Y (2010) Scalable influence maximization for prevalent viral marketing in large-scale social networks. In: KDD, pp 1029–1038

Chertkoff JM, Kushigian RH (1999) Don't Panic: The Psychology of Emergency Egress and Ingress. Praeger Publishers

Diamond J, McVay M, Zavala MW (2010) Quick, Safe, Secure: Addressing Human Behavior During Evacuations at LAX. Master's thesis, UCLA Dept of Public Policy

Dodds P, Watts DJ (2005) A generalized model of social and biological contagion. Journal of Theoretical Biology 232(4):587–604

Doherty W (1997) The Emotional Contagion Scale: A Measure of Individual Differences. Journal of Nonverbal Behavior 21(2)

Durupinar F (2010) From audiences to mobs: Crowd simulation with psychological factors. PhD dissertation, Bilkent University, Dept. Comp. Eng

Grandey AA (2000) Emotional regulation in the workplace: A new way to conceptualize emotional labor. Journal of Occupational Health Psychology 5(1):95–110

Hatfield E, Cacioppo JT, Rapson RL (1994) Emotional Contagion. Cambridge University Press

Helbing D, Farkas IJ, Vicsek T (2000) Simulating dynamical features of escape panic. Nature 407:487–490

JLBryan (2002) Behavioral response to fire and smoke. In: SFPE Handbook of Fire Protection Engineering, 3rd edn, National Fire Protection Association, pp 3315–3341

Jolliffe D, Farrington DP (2006) Development and validation of the basic empathy scale. Journal of Adolescence 29(4):589–611

Kempe D, Kleinberg JM, Tardos É (2003) Maximizing the spread of influence through a social network. In: KDD, pp 137–146

Kenny PG, Parsons TD, Gratch J, Rizzo AA (2008) Evaluation of Justina: A Virtual Patient with PTSD. In: IVA-08, Springer, pp 394–408

Kermack WO, McKendrick AG (1927) A contribution to the mathematical theory of epidemics. Proceedings of The Royal Society of London Series A, Containing Papers of A Mathematical and Physical Character (1905-1934) 115:700–721

Leskovec J, Krause A, Guestrin C, Faloutsos C, VanBriesen JM, Glance NS (2007) Cost-effective outbreak detection in networks. In: KDD, pp 420–429

Lundqvist LO (2008) Factor Structure of the Greek Version of the Emotional Contagion Scale and its Measurement Invariance Across Gender and Cultural Groups. Journal of Individual Differences 29(3):121–129

Mileti D, Sorensen J (1990) Communication of Emergency Public Warnings: A Social Science Perspective and State-of-the-Art Assessment

Murray JD (2002) Mathematical Biology, 3rd edn. Springer

Proulx G, Fahy RF (2008) Human behavior and evacuation movement in smoke. ASHRAE Transactions

Pugh SD (2001) Service with a smile: Emotional contagion in the service encounter. Academy of Management Journal 44(5):1018–1027

Rogers E (1962) Diffusion of Innovations. The Free Press, New York

Schelling TC (1973) Hockey helmets, concealed weapons, and daylight saving: a study of binary choices with externalities. Journal of Conflict Resolution 17:381–428

Smith CA, Ellsworth PC (1985) Patterns of cognitive appraisal in emotion. Journal of Personality and Social Psychology 4(48):813–838

Tsai J, Fridman N, Bowring E, Brown M, Epstein S, Kaminka G, Marsella S, Ogden A, Rika I, Sheel A, Taylor ME, Wang X, Zilka A, Tambe M (2011) ESCAPES: Evacuation Simulation with Children, Authorities, Parents, Emotions, and Social Comparison. In: AAMAS-11, ACM Press, pp 457–464