Computer Science and Information Systems 14(1):239–255 DOI: 10.2298/CSIS161015003G

Requirement Engineering Activities in Smart Environments for Large Facilities

Jorge J. Gomez-Sanz[1], Rafael Pax[1], Millán Arroyo[1], and Marlon Cárdenas Bonett[1]

Universidad Complutense de Madrid,
28040, Madrid, Spain
jjgomez,rpax@ucm.es, millan@cps.ucm.es,mcardenas@ucm.es

Abstract. Developing a large, but smart environment is a complex task that requires the collaboration of experts of different disciplines. How to successfully attain such collaboration is not a trivial matter. The paper illustrates the problem with a case study where the manager of the facility intends to influence pedestrians so that they choose a task that requires certain effort, e.g. using staircases, instead of the current one that requires less effort, e.g. using the elevator. Defining requirements for such scenarios requires a strong multidisciplinary collaboration which is not currently well supported. This paper contributes with an approach to provide non-technician experts with tools so that they can provide feedback on the requirements and verify them in a systematic way.

Keywords: crowd simulation, applied social sciences, ambient intelligence, multidisciplinary development.

1. Introduction

According to the International Facility Management Association [11], "Facility management is a profession that encompasses multiple disciplines to ensure functionality of the built environment by integrating people, place, process and technology". The efficiency of the use of the facility is one of the goals of a manager and for this aim, he/she must be willing to work proactively with new configurations of the facility [4]. For this aim, the incorporation of Internet of Things technology may enable a cheap enhancement of facilities that could be used to foster a more adequate use of the installation and a faster experimentation. This different use of the installation necessarily involves the alteration of the inhabitants of the facility, be them regular visitors or operators. The necessary technology and the possible, feasible, behavior alterations are questions whose answers require the collaboration of social scientists and engineers.

Basing on the hypothesis that people can modify their behaviors if the adequate stimuli are provided, a social scientist can identify a set of possible desired behaviors that could turn out. There are arguments supporting this hypothesis, as it will be discussed. Social marketing is an example of conduct alteration via visual, olfative, and auditive stimulation [15] towards the achievement of some social goal, such as improving healthy habits. However, the goal of this paper is to address the dynamic interaction with the inhabitants of a physical space.

Hence, the focus is on large groups of individuals whose mass behavior is to be subtly altered. Examples of subtle alterations may be, influencing the permanence in certain areas or preventing users to visit others. Such behavior alteration does not have to be permanent. A plan for stimulus intensity and typology is needed and means for enacting such plan provided.

Developing systems that realize this stimuli plan becomes troublesome because the gap from the requirements to the implementation is too huge. Hence, this paper illustrates ways in which this gap can be overcome using the idea of simulations as interlingua. These simulations serve to formalize the observations of the social scientists, but also to sustain the rest of the development providing a testing platform to evaluate the stimuli effect.

The use of simulations is not new in Social Sciences. However, the process by which they are created has not been studied thoroughly. Social scientists research tools are usually surveys and interviews. The way to proceed from these interviews to actual simulations is part of this contribution too. However, how these interviews and surveys translate into the system specification is not straightforward. With respect to this matter, the paper proposes using simple tools that enable to channel the effort of the social scientist into a specification workflow based on the extensive use of software simulations.

Preliminary work [9] has shown some guidelines and steps for the enactment of social sciences participation in a systems engineering process. Data and case study are reused here, but the focus is moved towards the workflow where the social scientist participates. In this sense, the original contribution are tools for retrieving feedback about the simulations and the notion of deliverable as a simulation plus comments as part of the process.

The paper is structured as follows. First, section 2 analyses if a particular behavior alteration/induction is really possible. Requirements engineering is the focus of the section 3. It proposes a set of activities to be performed within requirements elicitation, specification, and verification phases. Section 4 shows an example of how a requirements gathering is conducted in a case study of behavior modification. Section 5 introduces some of the issues found during this work. The related work is introduced in section 6. Conclusions are presented into section 7.

2. Relevant Stimulus Towards Behavior Modification

At this point, it seems necessary to address if a crowd behavior can be altered in the context of a large facility through the use of different stimuli. This argumentation is necessary step prior to engage into the analysis of the kind of behaviors that are to be enacted in a facility.

There are precedents that support such influence is possible. The first is the existence of Marketing discipline, where buyers are influenced to buy a specific item or service. This will be further discussed in the related work section (section 6).

In the conductivist research in social sciences and psychology it is possible to find evidences as well. The scope of intensity of the behavior alteration changes depending on the nature of the behavior to be modified. Subtle and small alterations, such as "staring" gesture, have been documented. If an individual finds a group along the way, depending on its size, she will either stop, look what is happening, and stay; or keep walking [19]. The larger the group, the greater the effect. This is explained as a mirroring behavior effect. If sufficient people stare at an arbitrary point, a passerby individual will unconsciously look at the same place [8]. Gaze copying happens mainly within 2 meters range and the response depends on the physical layout of the environment, the social context, and the sex of the individual.

The impact of controlling the ”staring“ gesture of crowds in a large facility is low, specially due to the cost of having hired people to initiate the effect. However, other reactions are more relevant because of the obtained effect and the cost of producing the stimulus through artificial sources. Sound and images can affect the behavior of pedestrians. Beyer et al. [2] introduce an experiment where an interactive large banner display affects the audience. Through visual stimulus, authors manage to attract approaching pedestrians and distribute them along the display. Miller [20] shows that noise can affect people’s performance. A sleepy person may be aroused by noise, but it has also negative effects, like affecting the performance of complicated tasks, affect negatively the mood and disturb relaxation. Negative effects could be used to influence pedestrians. In this paper, it is assumed that, since it can annoy people, this could be used to clear out areas or to reduce the pedestrian traffic around some places where the noise comes from.

With a more specific focus on promoting an effort taxing activity, Foster [7] reviews different works that address how the environment can promote health-enhancing physical activity (HEPA). Selected works have focused on the staircases vs elevator decision. The stimuli has been mainly posters with different kind of content and sizes and other visual aids, such as altering the aesthetics of the stairs. The effect has been to an average of a decrease of 3 points in the percentage of use of staircases. Among stimuli, videos and a combination of talking directly with individuals are not listed as used means according to [7].

Certainly, the attitude towards the stimulus depends also on the interest of the individual. The same individual may pay more or less attention to the same stimulation. A person in an airport will frequently check information panels, whereas in a mall center it may not be the case. Fun parks also influence the behavior of their visitors through information panels that tell expected waiting time for each attraction.

To understand the factors in the effectiveness of the stimulus vs the nature of the stimulus, the person, and the context, Wiebe’s analysis of influencing factors in four social marketing campaigns can provide clues, as [13] quotes. There are five factors: the force (predisposition towards the goal and the intensity of the message), the direction (knowledge of how or where the person might go to consummate his motivation), the mechanism (an agency that enables to translate the person’s motivation into action), the adequacy and compatibility (ability and effectiveness of the agency), the distance (the estimate of energy cost the person will invest to achieve the goal vs the reward).

Nevertheless, there are sufficient results that suggest that crowds can be influenced and its behavior modified. Humans are sensible to external stimulus. The outreach of the influence may depend on different factors, as it has been explained. In certain conditions, such as evacuations, humans pay more attention to other humans rather than other artificial elements, such as banners.

3. Requirements elicitation, specification, and verification activities

It is well known that failure in determining the requirements of a project is major cause of project failure [23]. In order to understand precisely how to achieve the kind of behavior a facility manager expects, simulations are needed and such simulations need to be validated.

To this purpose, a set of requirements elicitation actions are needed covering:

- The environment. The team has to understand the physical construction and what purposes each area serves to. Also the procedures that the facility has to pay special attention to, as the evacuation procedures. It is expected, for instance, that a cafetery gathers an important amount of traffic in a faculty building.
- Available stimuli. Literature has to be checked looking for related behaviors and associated stimuli, to evaluate the effect of each one and if there is already empirical evidence of the outcome. If the effect is not the one that permits to achieve the intended behavior alteration, a field study needs to be conducted evaluating different stimuli and their result.
- Default behavior. Documenting the initial behavior of the crowd is something very specific of social sciences. Classical methods such as recording videos, conducting surveys and interviews, can be used to gather information. A later analysis is needed anyway to evaluate the information and find clues of what stimuli could be more effective towards achieving the desired behavior.

The information captured at this stage is computationally represented through simulations by an engineer. This begins the requirements specification phase. However, the simulations have to deal with incomplete data. Tracking the movement of the individuals moving through a facility is not possible in general, due to budget or legal constraints[22]. Usually, the information about the movement of the people inside a building is given in form of datasets of pedestrian traffic measured on strategic checkpoints over time. Hence, it is a major issue how to imply or infer what behavior the untracked people had along the experiment. To address this issue, it is assumed that more than one simulation have to be produced, each one identifying pedestrian behaviors traits as precisely as possible. Following sections will introduce how an algorithm for population behavior is used to generate such initial simulation setups. These setups suggest different trajectories that satisfies the empirical data collected in the field experiment or available literature. Whether these simulations are realistic, in terms of expectations of human behaviors, is something that is evaluated later on by an expert.

If these computer simulations captured the reaction of the crowd when the relevant stimuli was produced, then they could be used as testbed for evaluating what stimuli sequence would be needed in each situation. To achieve this, basic reactions of the simulation actors need to be accurately represented so that, during the simulation and while producing the stimuli, the reaction of the crowd is a convincing one. To this aim, some basic reactive behaviors are coded in the simulated characters. Section 4.2 discusses this aspect.

The interplay between the engineers and experts is introduced in figure 1. Engineers produce the simulation that captures the expected behavior of the crowd,while the experts validate the simulations.

The simulation specification is performed by an engineer. The engineer will define a set of possible actor behaviors, an enumeration of the number of instances of these behaviors, and a timestamp of when the behavior & actor instantiation happens.This permits to define scenarios where simulated actors arrive to the facility at different times. Simulated actors will either be responsible of operating the facilities or just visitors. The first

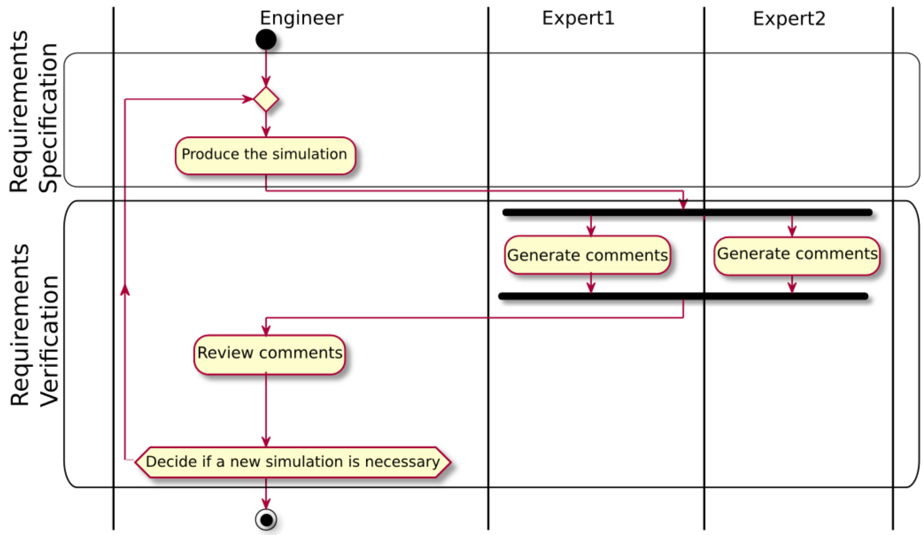

Fig. 1. Activity diagram showing specification and verification activities

group will be in charge of coordinating the behavior of the second group, just as a security guard guides people outside the building in an evacuation, for instance. Even though there maybe a role switch may happen in the literature, e.g. a citizen becomes a leader in an evacuation, it is not considered in this paper. Simulated actors conduct a specific predetermined behavior sequence due to scalability issues.

Each possible predetermined behavior represents a role of a person in the facility (e.g visitors, people working on the facility or security staff). They are called *predefined behaviors* because they define roughly the behavior of the role. Every actor behaving as a ”student” will behave in the same way as other students. However, future versions may consider small behavior variations.

As part of the requirements verification activities, it will be necessary to ensure the consistency and validity of the simulation. It is likely that many different simulations are produced as consequence of grounding variables such as the initial population or daily activities. The verification of all produced simulations can be achieved with the collaboration of the social scientists that produced the initial data. They need to observe the result of the computer simulation and subjectively decide if the simulation makes sense or not. According to figure 1, these experts are expected to produce approving/disapproving about the simulation.

Having an expert in sociology or social science can save the technician time and ensure an assessment of the model more in line with human behavior theories. For example, Boukas, Evangelos et al. [3] propose the same concept of social distance proposed in this paper, based on the field of proxemics, for an analog model of pedestrian traffic simulation. The collaboration of experts for the verification and validation of pedestrian traffic simulations is considered too in [6]. They suggest that the evaluation of the models requires a large amount of detailed data whose production consumes time and effort, making the identification and design of suitable criteria and data especially important. The involvement of these experts is not trivial either, as indicated in [5]. They provide an easy-to- use simulation environment for non-computer users for validation and verification. This type of environment that is easy to understand and use for any user would be the one that would give greater facilities so that the specialists of other fields outside the computational sciences to express and simulate to carry out the work of verification and validation of the model, as well as to propose improvements.

In any case, literature supports the hypothesis that experts are necessary, that they need to be involved to select the most relevant criteria and data, and that they need appropriate tools to do so. Though social scientists have been cited so far, other professionals could be added too. Psychologists as well as experts in safety and security can be included in the gathering and elicitation of requirements.

They need to observe the result of the computer simulation and subjectively decide if the simulation makes sense or not. Though social scientists have been cited so far, other professionals could be added too.

4. The case study

As a case study of facility management support scenario, two university faculties are considered: a computer science faculty and a social and political sciences faculty. In each faculty, it is intended to change the habit of using the elevator for a more healthy one which involves using more the staircases. Since the experiment deals with two different populations, it is an opportunity to study the reaction of two groups of people when facing the same stimuli. These two groups have relative homogeneity within each faculty, but there differences when comparing one to another.

The homogeneity of the two groups is given by the characteristics of its members, who are students, teachers and administrative staff of the same university. However, there are differences that need to measured: the daily activity of the faculties is different (e.g. schedules) and the attitude differs too when considering specific groups, such as students. The schedule, for instance, is more intensive in social and political sciences faculty from Monday to Thursday. Also, the building hosts an additional faculty, what increases the diversity in the schedules and the occupation of classrooms. The attitude differs too in the daily activities when considering students. As an example, in the social and political sciences faculty students more frequently express themselves in public through statements and public meetings. However, in the computer science faculty, the behavior of students is less reinvindicative. Besides, the student associations do involve themselves mainly into organizing computer related activities into closed rooms, rather than meeting into corridors or halls. The field study would point out quantitatively how these observed differences do affect the daily activities within these buildings.

An important part of the field study is the measurement of the effect of the stimuli in the crowd. Variables such as the already mentioned differences between the populations in the two buildings ought to affect the effectiveness of the stimuli. Through a phased field experiment, the effect of each stimuli had to be measured, so that it could be later on reproduced.

4.1. Requirements elicitation

Starting with the requirements elicitation activities stated in section 3, the environments were analyzed. In this case, both faculties were inspected looking for locations for running the behavior altering experiment. In this case, it required studying the behavior patterns of students and staff, looking for the elevators that were more used apparently, and that had staircases close by. For each location, observation points at different floors were identified too. Also, experimentation days were chosen taking into account when students and staff could have a peak in the occupation of the building.

The available stimuli in the literature was studied too. Some results have been cited here already in section 2 and others will be included in the related work section. Mostly, literature referred to banners situated along the path of pedestrians. For the experiment, three different stimuli are identified: banners containing some information that may affect the pedestrians; multimedia beamers showing videos motivating to show the staircase; and direct intervention. Different **banners** prototypes were made. Following the literature, it was intended to highlight facts that may be relevant to pedestrians, see figure 2. In summary, the banner suggests less time to reach the destination, more climate friendly activity, and health benefits because of the calorie burning. For the **multimedia**, a vertical beamer was used. This beamer projected over a panel with sufficient surface to attract the attention. It was the equivalent of a 55” screen. It played a video made for this occasion. It was a dramatization of a person that can walk, but does not want to, even wants to use a regular chair inside an elevator, and asks for people to raise the chair and get the person out. The **direct intervention** was made by volunteers that acted as if they were *facility operators* trying to influence a visitor. They stopped pedestrians with the excuse of doing a survey, but, in fact, they asked the users if they knew of the benefits of using staircases.

Fig. 2. Banner stimulus with information concerning the benefits of using staircases. It is written in Spanish. The main title says *stair climbing and avoiding elevators* at the top. The alleged reason are *1. improving your health*, *2. You will get faster to your destination*, and *3. you will save energy*

To evaluate the effect of the stimulus over the two faculties, a five week schedule was prepared:

- First week (week A). There was no stimulus and it was used to collect a base line of staircase/elevator traffic stats
- Second week (week B). The banner stimulus was introduced
- third week (week C1). The videos were added to the experiment
- Fourth week (week C2). Volunteers stop pedestrians to ask them if they know the benefits of going upstairs. To prevent rejection, the question is disguised as a survey.
- When week C2 finishes, all stimuli are removed from the environment.
- Fifth week (week D). This week happens a few days after week C2 finishes. The goal is to measure how much the behavior persist without stimuli reinforcement.

Along the experiment, a team of observers recorded in specific locations of the buildings how many people were traversing per minute the area and if they were going to get on or off the elevator, and if they were going upstairs or downstairs. By counting the people crossing these sections, the effect of the stimuli could be determined. Measurements and stimulation were made the same days each week, to replicate the same conditions as much as possible every time. The accumulated data, obtained from [9], is presented in table 1.

The number of people arriving through the elevator remains mostly the same along stages. However, the number of people choosing not to use the elevator when going upwards is reduced from 29.3 to 25.3 points in the percentage in phase C2. This variation is consistent with other experiments found in [7], where similar variations were found. In relative terms, the variation of 13,65% over the original use of the elevator. To evaluate the results, it should be taken into account that each stimulus lasted for one week, and not months.

Table 1. Variation of the traffic in elevators in two faculties as introduced in [9]

% use elevators	A	B	C1	C2	D
Total	23,1	21,9	21,4	20,3	22,2
Departures	29,3	28,2	26,2	25,3	26,4
Arrivals	14,4	14,4	15,2	14,0	16,2
#total=	9730	9797	9459	9165	9088
#departures=	5688	5371	5335	5109	5345
#arrivals=	4042	4426	4124	4056	3743

With the end of the field experiment and the analysis of the obtained data, the requirement elicitation finishes.

4.2. Requirements specification

The activities in requirements specification and requirements verification are expressed in figure 1. The subject of this section is the requirements specification part and will be repeated as the verification indicates the simulation or simulations are not correct.

As explained in section 3, the behavioral data collected by social scientists have to be translated into simulations that reproduce the observed behavior. The whole population movement is unknown and if only the information gathered during the field study is used, there is a potential unlimited number of simulations whose behavior matches the data. For instance, if the observed data shows a section was crossed twice, a simulation that would fit the data could be one where there is a single individual crossing twice that section; but another possible simulation would have two individuals crossing once the same section. Hence, the simulation specification poses a strong parameterization problem.

Actors in the simulation are playing the visitor role and there is none with *facilities operator*. Their basic behavior consists of entering the building, visiting some rooms, and then getting out of the building. During parameterization, a number of visitors is chosen, and for each visitor, the number of rooms to be visited and the sequence, and how much time will be invested in each location.

Actors move across the terrain facing collisions with obstacles and among each other. The specific path, speed, and delays that the actor finds adds some uncertainty to the result. There is no coordination in the current version of the simulation, so actors do not agree on the path. However, there is some implicit agreement when considering imminent collisions among two individuals. There is a path replanning to avoid the collision, which is close to what humans do.

Under these constraints, a first goal is, given a predetermined population size, find what behavior has to be assigned to each individual so that the execution of the resulting navigation path provides the same traffic data as observed. Once this goal is achieved, the next step is to reproduce the stimulus effect on the individual actors to obtain variations in the observed traffic similar to the experimentation, i.e., a 4 percentage points of variation of the traffic at the observed locations.

At the end of the process, the resulting traffic in all stimulus scenarios ought to match those of table 1. Achieving this turns out to be challenging because of the unexpected collisions, accumulation of individuals in crowded places, or the integration of elevators in the problem. About the later ones, elevators, see figure 3, are one of the most difficult problems to model. They require the simulated actors to coordinate the arrival and the evacuation of the elevator. If the elevator stops at the first floor, the simulated actor wants to get to the second, but the actor is at the front part, it should either get out or move aside to let others exit first.

Even if there is a simulation that matches the observed field study data, from the perspective of the social scientist, it may not make sense. As an example, the movement of characters may look unnatural if all characters decide to turn around in the middle of a corridor. That is the reason why, even though a simulation is data-wise correct, it may not be the one the project pursues. This problem has been studied in [21] and an algorithm proposed that generates a simple population.

To observe the simulation, several cameras are needed. Figure 4 shows an evolution of the work done in [9]. On the left hand side, there are five selectable cameras that show different locations of a building. On the main screen (center), the principal traffic location is shown, which matches the main elevators and staircases. The building matches the faculty of computer science in this case. The design of the simulation is such that the cameras, in this case, are pointed out to the places where observation data was collected. The bottom part of the interface shows the recorded data against the real time generated

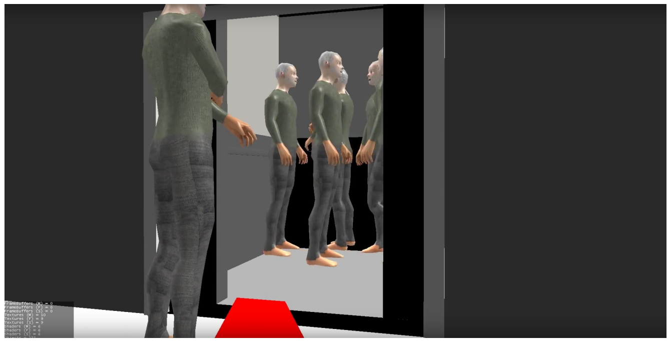

Fig. 3. A capture of the incoming and outgoing traffic to the elevator

traffic data. Ideally, both lines (generated and recorded data) ought to be close, but there is a variation. The reason is the above mentioned unexpected collisions and bottlenecks.

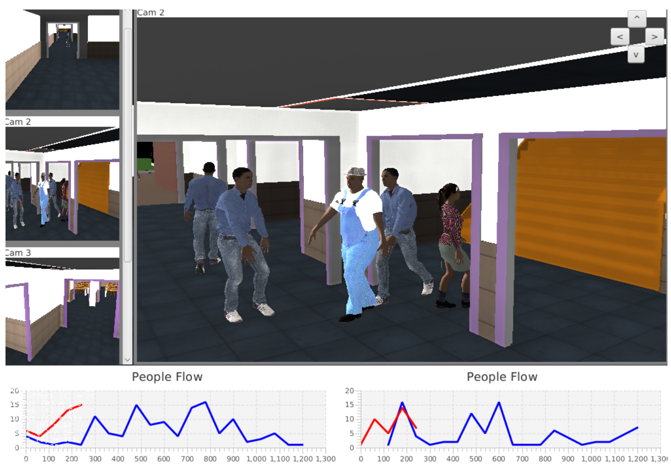

Fig. 4. User interface of the simulator, showing different cameras that permit to observe the behavior of people and charts to summarize quantitative data on some events. There are two lines, a complete line in blue representing the measured pedestrian traffic in the real world, and another growing line in red representing the measured traffic along the simulation.

Stimuli and simulations An important problem consists in quantifying the reaction of the people to the different stimuli, as measured by the field study. The simulations, so far, reproduce separately specific phases of the experiment, phases A to C2 according to table 1. Hence, the developer has to produce as many simulations as phases considered along the experiment. In each simulation, the goal is to reproduce the empirical data through a simulation.

A later step consists in experimenting with the simulations alone, altering a simulation corresponding to one field experiment phase to include stimuli considered into another phase. This is a more complex step not considered in this paper.

The course of action would involve including reactive behaviors into each simulated actor. Such individual reaction is already coded in the current version. For instance, this enables a simulated character to avoid collisions with other characters, or to alter the navigation when some unexpected bottleneck arises in a corridor.

4.3. Requirements verification

The activities in requirements verification included in figure 1 deal with the analysis of the simulations to determine whether they are correct or not.

The obtained simulations, as those from figure 4, need to be evaluated by the experts, social scientists in this case. For this goal, a way for delivering the simulations to the experts and get feedback was necessary. Figure 5 introduces the tool. The simulation is re-run to obtain a separated video per camera. These videos are integrated in an HTML5 and JavaScript application. The reason for an HTML5 and JavaScript implementation has the advantage of working in almost any computer as long as there is a browser.

The expert has to review the different videos which are played in a synchronized way. The review is aided with controls for freely moving across the videos: a slider control for jumping to specific sections; play and pause buttons; a framerate modification button; and controls for stepping forwards/backwards a few frames.

The expert has to visually inspect the simulation and then annotate in each captured camera whatever information to be recorded as feedback. Annotations are associated to the current timestamp and are specific of a camera feed. The expert can select an annotation and make the simulation move to the associated timestamp, what implies moving forwards or backwards all videos. To facilitate the later review of the comments, the expert can add a flag to the annotation to express whether the annotation is positive (the simulation is correct at that point) or negative (the simulation is incorrect). The positive annotation is not required every time and it is intended for informing that a failure in a previously reviewed simulation has been fixed. It means that the problem has been solved. Also, it is useful to point at behaviors that have to be preserved till the end of the development of the simulation, like a regression test.

Once the video is annotated, the expert can store the changes and send them back to the team. Then, the annotations can be checked looking for negative annotations. These are reviewed and included in the documentation of the development as a *deliverable*.

For example: In the context of a simulation about daily activities in a building, the behavior of the simulated users, at certain simulation time, might not match the expectations of the social sciences expert, such as having people insisting in getting closer to a crowded exit. The social science expert would address this issue through the annotation tool described previously. This expert would add a negative comment pointing out the irrationality of such behavior. Later on, a review of the comments of the expert would aid the development team to identify which individual or group behaviors are incorrect. In this case, the solution is to make simulated users avoid getting stuck into crowded exits. Then, the development team would produce a new simulation addressing these issues.

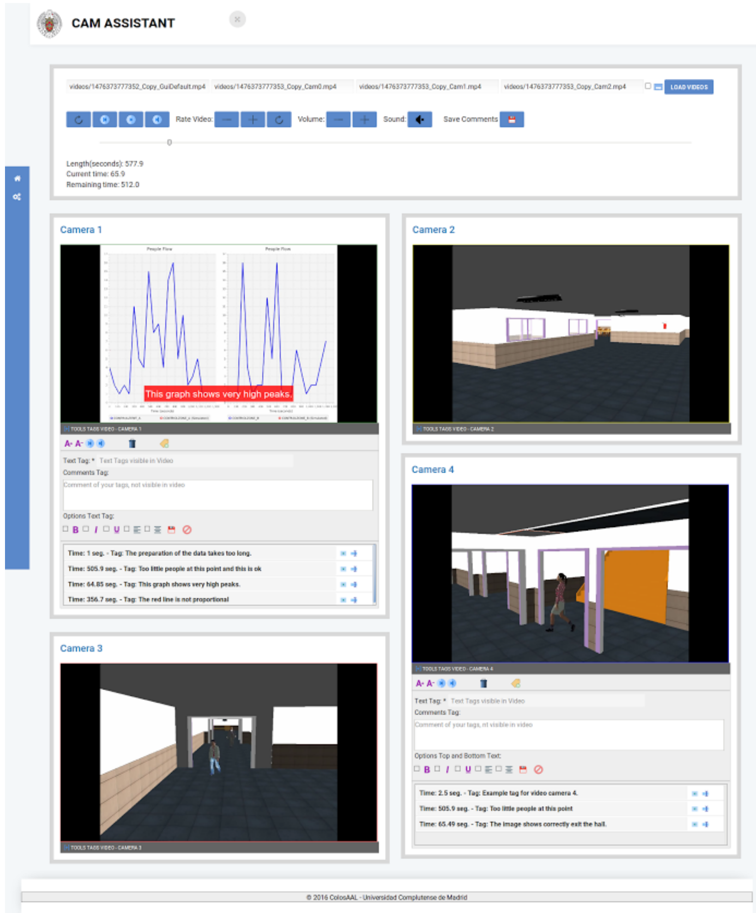

Fig. 5. Annotation tool for adding comments to the simulation

The expert would review again and convert the negative comment into a positive one. Along the process, the produced artifacts would be an initial simulation, some comments (positive and negative) generated by an expert, and a new simulation that addresses the comments.

This manual validation made by experts is necessary because the simulation may contain inconsistent results. Human experts can notice these inconsistencies and decide if they really happen in the real world or not. As an example, a simulation may have most simulated actors gathered at the lower floors and having little or no people at the upper ones [21]. In the real building, top floors do not have classrooms, only offices. If top floors really were empty, the facility manager may consider if the building was being effectively used or not. However, the human expert may think too that the simulation is wrong and that the building cannot have empty top floors. In this case, top floors are almost empty and have less pedestrian traffic than lower floors.

5. Evaluation

Generating an annotated version of the simulation and defining the simulation itself is proving to be an effort taxing task, mainly because of the inherent technological problems when defining the behavior of the crowd. It is not the definition of the population, but the management of the unexpected events such as collisions and bottlenecks what is adding cost to the development.

The work done has not taken into account stakeholders, such as the *facility manager* or the staff in the facility. However, it is expected that the simplicity of the process in annotating videos makes the task affordable in terms of human-to-computer interaction. The feedback mechanism is the same despite the background of the expert. Only the content of the annotations will differ, which is something expected due to the different perspectives of each reviewer.

The availability of different cameras makes the review simpler. In the GUI, see figure 4, the user can switch cameras by selecting one on the left. The annotation tool, see figure 5, is more flexible, since the user can reorganize the videos and even make the window larger to see them all at once. However, it will require a larger monitor.

Processing these annotations is another time consuming task. It is not only about reading the comments, but also classifying them and making sure that new simulations do not fall into the mistakes while still maintaining the behavior the positive annotations identified. By now, this needs to be done manually, but it has pointed out a non trivial problem, which is how to formally capture the comments so that they can be automatically verified across simulations.

About the behavior of the simulated crowds and the field experiment data from figure 4, there is still not a perfect match. Improvements made to work in [9], such as having different people sizes (people in the figure 4 is shorter than people in figure 3) has created greater uncertainty in the outcome of the simulation since the collisions are more frequent. This implies simulated characters take longer to achieve the destination, so they cross the control sections at a different time. This naturally affects the final accounting of simulated pedestrian traffic. This problem will become harder after adding the social concerns that the simulation must include, such as constraining the movement of individuals to fit into social-etiquette conventions. For instance, if the place is not crowded, two persons will be close if they know already each other, or one wants to address the other. Otherwise, both will keep the distance. Other relevant concerns are moving in groups or pairs, avoiding bottlenecks, or dealing with rooms' exits in an ordered way.

6. Related work

There are works dealing with the design of smart systems, but they do not frequently consider human sciences and stimulus to plan the kind of system which is needed and what performance it will have. Harrison [10] claims the analysis of mutual and incidental user interaction has not been accounted and proceeds to apply fluid flow analysis to understand it. This kind of analysis is necessary, but, it does not replace a more conventional study and cannot assume a 100% response of the individuals every time. Other works focus on the devices expected to provide the stimulus at small scale, such as [24]. Though authors stress the involvement of human scientists too, the behavior of people in small spaces cannot be compared to that of large spaces.

There are precedents too in reproducing observed data as simulations. In [14], video recordings were used to reproduce later on a crowd simulation of simulated actors. Behavior of the individuals were obtained from a multiple checkpoint observation that allowed to reproduce the pedestrian traffic of the facilities where the measurements were made.

The project introduced in this paper, however, assumes incomplete information about activities and traffic. The less information is used, the less expensive a real installation would be. Following the same paradigm, Lerner et al. [16] propose the creation of an example database for evaluating simulated crowds based on videos of real crowds. Bera et al. [1] also developed a behavior-learning algorithm for data-driven crowd simulation, capable of learning from mixed videos. Zong et al. [25] developed a framework for generating crowds for matching the patterns observed on video data,taking into consideration the behavior both at the microscopic level as at the macroscopic level. Finally, Yi Li et al. [17] developed a technique for populating large environments with virtual characters, cloning the trajectories of extracted crowd motion of real data sets to a large number of entities.

One discipline that has extensively researched the behavior of people while wandering through large commercial facilities is Marketing. Marketing is about the exchange process, “where two or more parties, each having something to exchange, and both able to carry out communications and distribution” [13]. Then, marketing management could be defined as “the analysis, planning, implementation, and control of programs designed to bring about desired exchanges with target audiences and the purpose of personal or mutual interest” [13]. In a way, Marketing wants to modify a behavior, a buying behavior, of a target group of people, the target customers. The approach could be applied to other less business-like goals, such as increasing charity donations or teaching cultural institutions how to attract new sponsors. It is the social marketing [13][12], and pursues “the design, implementation and control of programs calculated to influence the acceptability of social ideas“. Though this goal differs from the project, there are points in common in the use of stimulus to alter the behaviors. More recently, McKenzie-Mohr [18] uses marketing to raise concern on climate change and induce less contaminating behaviors. Marketing provides valuable lessons in how to analyze and handle the stimuli, though from the perspective of business orientation, classic Marketing, or a short term controlled influence, like social marketing. The work introduced in this paper is more related to long term stimuli production and a dynamic configuration of the stimuli to adequate the behavior of the crowd according to the requirements expressed by stakeholders such as a *facility manager*.

7. Conclusions

The paper has worked on specific phases of Requirements Engineering to clarify the role of social scientists in the development and introduced tools that help to achieve the goal of formally documenting with simulations a crowd behavior alteration scenario.

Previous work [9] was oriented towards incorporating simulations into a larger development guideline. This paper has contributed with a special focus into requirements engineering activities, in particular requirements specification and requirements validation.

For this aim, a set of activities has been suggested that depends on a specific annotation tool introduced in this paper.

The collaboration of experts, in this paper, social scientists have mentioned, is a critical element. Since the project aims to capture the reaction of the crowd towards stimuli, an expert needs to assess that the simulation is actually showing a similar reaction. Also, the expert has to approve that a simulation is actually representing a real behavior. In this work this is even more important, since the simulation is built using partial pedestrian traffic data. Experts need to evaluate whether the recreation produced by the engineer actually resembles a real behavior.

Correctly capturing this feedback provided by the collaborating experts is an important step in order to guarantee the quality of the final development. Since the requirements engineering activities do extensively use simulations as formal specification tool, this paper has proposed to understand this feedback process in form of annotations made to a video that represents the simulation of a particular scenario. There are several advantages for this approach: the problem is reduced to something that requires no additional means aside a browser; the actions required on behalf the expert limit to annotating a video, something that should be intuitive enough; and the involved artifacts, videos and comments, can be easily stored and reproduced anytime, also they can be subject of configuration management activities, such as controlling their evolution through version control tools.

However, the generation of simulations itself remains a specialized activity that still requires the participation of engineers. It remains as future work to develop means that facilitate the online creation of these simulations and, perhaps, the collaboration among experts to create them.

In the paper, simulations reproduce empirical pedestrian traffic of specific field experiment phases. Hence, there is at least one simulation per field experiment phase. It remains as future work to generate simulations where an operator can tune the behavior in runtime, add the stimuli to the simulation, and observe how the crowd behavior changes to fit experimental data.

Another challenge is how to systematically deal with sets of annotations produced by different experts. A development team may not have the criteria to decide, when conflicting comments arise, which one should be the more chosen in the new simulation. It is expected that techniques like focus group and other qualitative analysis techniques help to organize and prioritize the different comments.

Acknowledgments. We acknowledge support from the project "Collaborative Ambient Assisted Living Design (ColoSAAL)" (TIN2014-57028-R) funded by Spanish Ministry for Economy and Competitiveness; and MOSI-AGIL-CM (S2013/ICE-3019) co-funded by Madrid Government, EU Structural Funds FSE, and FEDER.

References

1. Bera, A., Kim, S., Manocha, D.: Efficient trajectory extraction and parameter learning for data-driven crowd simulation. In: Proceedings of the 41st Graphics Interface Conference. pp. 65–72. Canadian Information Processing Society (2015)

2. Beyer, G., Binder, V., Jäger, N., Butz, A.: The puppeteer display: attracting and actively shaping the audience with an interactive public banner display. In: Proceedings of the 2014 conference on Designing interactive systems. pp. 935–944. ACM (2014)
3. Boukas, E., Crociani, L., Manzoni, S., Vizzari, G., Gasteratos, A., Sirakoulis, G.C.: An intelligent tool for the automated evaluation of pedestrian simulation. In: Hellenic Conference on Artificial Intelligence. pp. 136–149. Springer (2014)
4. Cotts, D.G., Roper, K.O., Payant, R.P.: The facility management handbook. AMACOM (2009)
5. Federici, M.L., Manenti, L., Manzoni, S.: A checklist for the evaluation of pedestrian simulation software functionalities. arXiv preprint arXiv:1404.7717 (2014)
6. Fetiarison, M., Flötteröd, G., Bierlaire, M.: Evaluation of pedestrian data collection methods within a simulation framework. In: AET, European Transport Conference. Glasgow, Scotland. 11/10/2010–13/10/2010 (2010)
7. Foster, C., Hillsdon, M.: Changing the environment to promote health-enhancing physical activity. Journal of sports sciences 22(8), 755–769 (2004)
8. Gallup, A.C., Hale, J.J., Sumpter, D.J., Garnier, S., Kacelnik, A., Krebs, J.R., Couzin, I.D.: Visual attention and the acquisition of information in human crowds. Proceedings of the National Academy of Sciences 109(19), 7245–7250 (2012)
9. Gomez-Sanz, J.J., Pax, R., Arroyo, M.: Towards the simulation of large environments. In: 1st International Workshop on Ambient Intelligence for Large Premises (AmILP 2016). CEUR Proceedings (2016)
10. Harrison, M.D., Massink, M., Latella, D.: Engineering crowd interaction within smart environments. In: Proceedings of the 1st ACM SIGCHI symposium on Engineering interactive computing systems. pp. 117–122. ACM (2009)
11. International Facility Management Association: (October 2016), https://www.ifma.org/about/what-is-facility-management
12. Kotler, P., Roberto, E.L.: Social marketing. strategies for changing public behavior. (1989)
13. Kotler, P., Zaltman, G.: Social marketing: an approach to planned social change. The Journal of Marketing pp. 3–12 (1971)
14. Lee, K.H., Choi, M.G., Hong, Q., Lee, J.: Group behavior from video: A data-driven approach to crowd simulation. In: Proceedings of the 2007 ACM SIGGRAPH/Eurographics Symposium on Computer Animation. pp. 109–118. SCA '07, Eurographics Association, Aire-la-Ville, Switzerland, Switzerland (2007)
15. Lee, N.R., Kotler, P.: Social marketing: Influencing behaviors for good. Sage Publications Inc (2011)
16. Lerner, A., Chrysanthou, Y., Shamir, A., Cohen-Or, D.: Data driven evaluation of crowds. In: Motion in Games, pp. 75–83. Springer (2009)
17. Li, Y., Christie, M., Siret, O., Kulpa, R., Pettré, J.: Cloning crowd motions. In: Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation. pp. 201–210. Eurographics Association (2012)
18. McKenzie-Mohr, D.: Fostering sustainable behavior: An introduction to community-based social marketing. New society publishers (2013)
19. Milgram, S., Bickman, L., Berkowitz, L.: Note on the drawing power of crowds of different size. Journal of personality and social psychology 13(2), 79 (1969)
20. Miller, J.D.: Effects of noise on people. The Journal of the Acoustical Society of America 56(3), 729–764 (1974)
21. Pax, R., Gomez-Sanz, J.J.: A greedy algorithm for reproducing crowds. In: Trends in Practical Applications of Scalable Multi-Agent Systems, the PAAMS Collection, 14th International Conference, PAAMS 2016, Sevilla, Spain, June 1-3, 2016, Special Sessions. pp. 287–296 (2016)
22. Rajpoot, Q.M., Jensen, C.D.: Video surveillance: Privacy issues and legal compliance. Promoting Social Change and Democracy Through Information Technology p. 69 (2015)
23. Schmidt, R., Lyytinen, K., Mark Keil, P.C.: Identifying software project risks: An international delphi study. Journal of management information systems 17(4), 5–36 (2001)

24. Streitz, N.A., Röcker, C., Prante, T., Van Alphen, D., Stenzel, R., Magerkurth, C.: Designing smart artifacts for smart environments. Computer 38(3), 41–49 (2005)
25. Zhong, J., Cai, W., Luo, L., Yin, H.: Learning behavior patterns from video: A data-driven framework for agent-based crowd modeling. In: Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems. pp. 801–809. International Foundation for Autonomous Agents and Multiagent Systems (2015)

Jorge J. Gomez-Sanz is Associate Professor in the Universidad Complutense de Madrid. He is one of the heads of the GRASIA research group. He has a degree in software engineering and a PhD in intelligent systems. His current research lines are ambient intelligence and model driven development.

Rafael Pax is researcher in the Universidad Complutense de Madrid. He has a degree in software engineering. He is doing a PhD in crowd simulation. He is a member of the GRASIA research group.

Millán Arroyo Menéndez is Associate Professor in the Universidad Complutense de Madrid. He has a degree and a PhD in political sciences and sociology. He is currently involved in the definition and integration of social concerns into ambient intelligence developments. He is a member of the GRASIA research group.

Marlon Cárdenas Bonett is researcher in the Universidad Complutense de Madrid. He has a degree in software engineering. He is currently doing a PhD in methods for the development of ambient intelligence systems. He is a member of the GRASIA research group.

Received: October 15, 2016; Accepted: January 10, 2017.