The Creation of Simulated Activity Datasets Using a Graphical Intelligent Environment Simulation Tool

J. Synnott, L. Chen, *Member, IEEE*, C.D. Nugent, *Member, IEEE* and G. Moore, *Member, IEEE*

Abstract— **The availability of datasets capturing the performance of activities of daily living is limited by difficulties associated with the collection of such data. Software solutions can mitigate these limitations, providing researchers with the ability to rapidly generate simulated data. This paper describes the use of IE Sim to create a simulated intelligent environment within which activities of daily living can be performed using a virtual avatar. IE Sim has been demonstrated to facilitate the generation of datasets capturing normal activity performance in addition to overlapping activities and abnormal activities such as hazardous scenarios.**

I. Introduction

There is currently a global trend in population ageing due to reductions in fertility and mortality combined with increases in life expectancy [1]. The result of this is an increasing demand on healthcare resources as the older population is subject to higher rates of cognitive and physical impairment compared with the younger population [2].

Dementia is one example of a cognitive impairment that increases in prevalence with age [3]. It is a progressive condition with several cognitive symptoms including memory loss, difficulty with planning or concentration and orientation in terms of time and location. As the condition progresses, the symptoms can affect the ability to complete everyday tasks including activities of daily living (ADLs) and instrumental ADLs (IADLs). [4].

Intelligent environments (IEs) are one solution that may alleviate the demand on healthcare resources. IEs are sensorized environments that are capable of monitoring the state of the environment and its inhabitants. Through the analysis of datasets generated within IEs, it is possible to infer the activities performed by inhabitants and derive metrics describing the performance. Such metrics may describe elements such as the time, location, duration and frequency of activity performance, in addition to the presence and order of individual steps performed within the activities themselves.

Invest Northern Ireland is acknowledged for supporting this project under the Competence Centre Programme Grant RD0513853 - Connected Health Innovation Centre. Additional support provided by the Department of Employment and Learning, Northern Ireland.

J. Synnott, C.D. Nugent and G. Moore are with the Computer Science Research Institute and the School of Computing and Mathematics, University of Ulster (phone: +44 2890 368840 e-mail: {j.synnott, cd.nugent, g.moore} @ulster.ac.uk).

L. Chen is with the School of Computer Science and Informatics, De Montfort University (e-mail: liming.chen@dmu.ac.uk)

The remote monitoring of these activities and metrics can allow those suffering from chronic illness to continue living independently in their own homes for longer than is possible with existing healthcare practice and may improve clinical effectiveness through decreased numbers of hospital admissions and reduced average length of stay. This is due to both the real-time and long-term benefits of remote activity monitoring. Real-time activity monitoring aims to facilitate the detection of events that may indicate a danger to inhabitant health. This may include areas such as fall detection or the detection of hazardous situations as a result of an activity incompletion, such as an appliance being left switched on. Other areas include security issues such as a door being left open at night, late night wandering, or issues with medication adherence. Long-term benefits of activity monitoring may facilitate earlier detection of cognitive or physical decline, presenting the opportunity for earlier intervention and maximization of quality of life. The monitoring of ADLs is viewed by medical professionals as being one of the most effective methods of detecting emerging medical conditions [5].

The development and testing of novel data driven approaches for activity monitoring relies on the availability of datasets containing activity data. Nevertheless, the availability of such datasets is limited due to a number of issues associated with the acquisition of such data. This paper discusses the use of IE Sim, a graphical IE simulation tool that facilitates creation of simulated ADL datasets through interaction with simulated environments. Section 2 provides an overview of related research, discussing the issues associated with the collection of activity datasets and describes popular techniques for the generation of simulated ADL datasets. Section 3 describes IE Sim in further detail, the simulated activities and the method by which these activities are recorded using IE Sim. Section 4 provides a discussion of the recording process and resulting datasets, and Section 5 provides concluding remarks.

II. Related Research

IEs are time consuming to construct and expensive to implement due to cost of the construction of physical environments and the sensor technology to be placed within them [6]. Additionally, the optimal layout of sensors within the IE may not be known prior to construction and physical environments are limited in terms of flexibility and scalability [7]. The range of scenarios that can be tested may also be limited for a number of reasons. Recruitment of suitable participants to generate data containing all scenarios may prove difficult [6], and it may also not be possible to record all possible scenarios as certain situations may be unethical to test on vulnerable patients [7]. There are also regulatory limitations that must be adhered to during testing on human subjects [6]. The difficulty in the collection of IE sensor data can therefore be viewed as being detrimental to the progress of research and is slowing advances in the development of new approaches [8]. The generation of simulated activity datasets can mitigate these limitations and provide researchers with sufficient data to aid the development and testing of novel data driven approaches to activity analysis.

One popular approach to IE data simulation involves the use of parameterized data models. Parameterized approaches for IE data simulation involve the specification of activity models that define the presence and order of events, the probability of events occurring and the time taken for each event during the performance of specific activities. Mendez-Vazquez *et al.* [9] demonstrated the use of Markov chains describing the order of events, combined with Poisson distribution to calculate a range of realistic activity times and probability distributions to calculate a range of sensor values to generate a simulated activity dataset. This dataset contained a distribution of activities such as reading, sleeping, walking and sitting, together with metrics including time and energy expenditure. Noury and Hadidi [10] developed a simulator utilizing Hidden Markov Models based on passive infrared (PIR) data collected within hospital rooms and apartments. This approach facilitated the generation of simulated datasets based on real data, providing good correlations with the real dataset.

Parameterized approaches have the potential to generate extensive synthetic datasets describing activity performances over extended periods of time. The quality and accuracy of the resulting datasets, however, relies heavily on the quality of the activity description model and associated parameters. The construction of accurate activity models requires access to real test data describing the performance of the modelled activities. Additionally, it may be difficult to accurately and intuitively adjust such models to represent subtle yet significant differences in activity performance. For example, the impact of a phone ringing in a living room area during completion of the "making a cup of tea" activity in the kitchen, or the impact of adjustments to the environment and sensor layout on the quality of data generated.

The use of interactive simulated IEs combined with embedded avatars has the potential to provide an intuitive and interactive environment simulation experience. Avatars are interactive objects that can move within simulated IEs and passively or actively interact with the objects contained within them, representing the behavior of real inhabitants within physical IEs. The avatar-based approaches rely on the modelling of environments and individual sensors contained within them. Such models may be based on existing environments or sensor specifications, or may be entirely conceptual. These approaches have the advantage of facilitating investigations into the impact of environment layout and sensor placement on the quality of the resulting dataset to be undertaken. Helal *et al.* [11] introduced Persim 3D, a context driven approach that incorporated activity scheduling to drive actions performed by a virtual character. These actions generated sensor data in the Sensory Dataset Description Language format. The proposed approach included the creation of 3D environments populated with objects and sensors, facilitating the visualization and user inspection of simulation progress.

Approaches that facilitate direct avatar interaction with simulated IEs enable the recording of natural activity performance. Activities can be performed in a natural manner by movement of the avatar within the environment and interaction with objects contained within the environment. This facilitates ad-hoc testing through the recording of specific activity scenarios with complex variations, such as interruption during activity performance, the performance of multiple specific overlapping activities simultaneously or the impact of subtle changes in object and sensor placement on the data generated. Krzyska [12] developed a simulation tool that facilitated the creation of simulated IEs containing motion sensors with adjustable sensing radius. An avatar could be moved within the environment using mouse clicks, which generated sensor events within a log file if the avatar moved within a movement sensor's detection radius.

There is the need for a flexible and customizable approach to IE data synthesis that facilitates the capture of complex ADLs. Such an approach must provide control over the environment in which they are performed, the properties of the objects involved, and the sequence of events. The following sections describe such an approach in detail, providing examples of its usage.

III. Proposed Approach For the Generation of Simulated Datasets

IE Sim aims to expand upon existing research in the area by providing a flexible yet intuitive interface for the generation of simulated datasets with direct avatar interaction. This software facilitates the creation of 2D overhead plans representing simulated IEs containing objects and sensors. Interaction with these IEs is performed using an interactive avatar and interaction menu. A user can navigate the avatar throughout the IE using the arrow keys on a keyboard, passively interacting with sensors or actively interacting with objects using the interaction menu. The result of these interactions can be adjusted by setting individual sensor properties for values including interaction range, delay, output values and interaction method. The interactions generate simulated sensor datasets in the homeML format [13]. IE Sim has received positive feedback from the research community during a usability study, rating its functionality in terms of ease of use and usefulness in research. A full overview of the technical implementation of IE Sim and the results of the usability study are provided in [14].

In this paper, the utility of IE Sim to generate datasets describing the performance of ADLs is demonstrated. The ADL of the preparation of a cup of tea is used as an example. This ADL was chosen due to its common nature in addition to its inherent complexity due to the number of steps involved, combined with the variety of valid and invalid combinations of steps that may be performed. This ADL may be performed in a vast number of different ways due to personal preference and overlap with other activities. A constrained ADL graph (Fig. 1) was created to guide and limit the number of variations in activity performance to be recorded in the current work. This graph represents a valid set of mandatory and optional steps involved in the ADL. This example provided a constrained subset of 8 valid combinations of steps. In addition to these valid step combinations, IE Sim was used to generate comprehensive datasets involving overlapping activities and potentially hazardous scenarios.

A. Environment Creation

In order to record the ADLs, a simulated IE was created using a selection of objects from the object toolbox included within IE Sim. This simulated IE was designed to contain rooms and objects found in most single-floor home environments. The 2D floor plan of the simulated IE is illustrated in Fig. 2. The IE consisted of a kitchen (Fig. 2 (D)), living room (Fig. 2 (E)), bathroom (Fig. 2 (H)), hallway (Fig. 2 (I)) and bedroom (Fig. 2 (K)).

The objects used to create the IE included rooms, a range of typical static items of furniture, and several passive and active sensors. The passive sensors included PIR sensors (Fig. 2 (C)), pressure sensors (Fig. 2 (F)), temperature sensors (Fig. 2 (B)) and humidity sensors (Fig. 2 (A)). These sensors generate data in response to sensor settings and the movement of the avatar within the IE. Active sensors are attached to objects that require direct user interaction using the simulation interaction toolbox, which is a context menu that provides interaction options based on object state and avatar proximity. These included door contact sensors in hallway doors (Fig. 2 (G)) and cupboards, a telephone (Fig. 2 (J)) and a lamp (Fig. 2 (L)).

B. Object Creation

IE Sim provides a number of default objects and sensors for inclusion within IEs. Nevertheless, for specific use cases, additional objects and sensors may be required. IE Sim supports the creation of custom sensorized objects using a built in user interface. Several such objects were required in order to complete the recording of the “make a cup of tea” ADL. Each of these objects was designed to simulate an object with a contact sensor or toggle switch attached. The interaction method, appearance, range, frequency of data generation and data values output by sensors can be configured during creation using the IE Sim user interface.

The ability to create additional object types results in a solution that can be applied to a wide variety of IEs. In

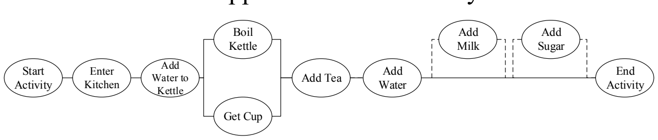

Figure 1. A constrained ADL graph illustrating an example of the “making a cup of tea” activity. Step branches represent mandatory steps that may be completed in any order before beginning the next step. Steps with dashed lines before and after represent optional steps.

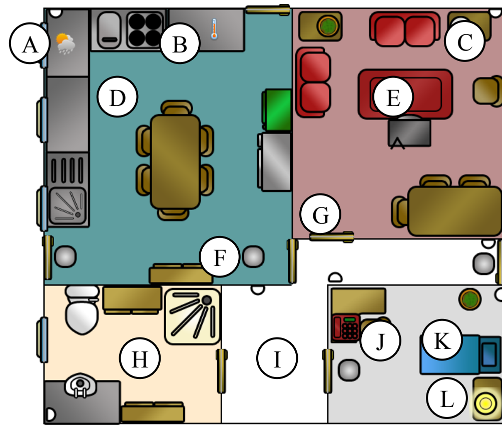

Figure 2. The simulated IE designed to represent a home environment. (A) Humidity Sensor; (B) Temperature Sensor; (C) PIR Sensor; (D) Kitchen; (E) Living Room; (F) Pressure Sensor; (G) Door Contact Sensor; (H) Bathroom; (I) Hallway; (J) Telephone; (K) Bedroom; (L) Lamp

addition to home environments, these may include environments such as office spaces, residential care facilities, hospitals and public spaces.

IV. Results

A dataset describing performances of the ADL was generated once the setup of the simulated IE was complete. IE Sim generates simulated sensor data once the avatar is placed within the created environment. Fig. 3 provides an example of an avatar (Fig. 3 (B)) that has been placed within the bedroom and positioned near the door. The PIR sensor (Fig. 3 (C)) detects the avatar’s presence within the room and the pressure sensor (Fig. 3 (A)) detects contact with the avatar. Detection of avatar presence is indicated by a red glow on each of these sensors. The simulation toolbox (Fig. 3 (D)) provides a menu listing the interaction options that are within range of the avatar.

The ADL performances were recorded by navigating the avatar throughout the IE and completing the valid combinations of steps as outlined in the constrained ADL graph. Fig. 4 displays the avatar in the kitchen together with the custom objects created for use in the performance of the ADL. Several additional scenarios were recorded in order to demonstrate IE Sim’s ability to generate diverse and comprehensive datasets. These examples demonstrated activity overlap and hazardous scenarios. The example of activity overlap involved the avatar beginning the “making a cup of tea” activity by switching the kettle on. While the kettle was on, the avatar navigated out of the kitchen and into the living room to switch on the TV and sit on the sofa. This generated multiple door contact sensor events, PIR activation events, a contact switch event on the TV and pressure sensor events from the sofa. After remaining in the living room for several minutes, the avatar then returned to the kitchen to complete the rest of the activity.

Three examples of potentially hazardous events were recorded. These include activities that indicate abnormal behavior that may require intervention. These examples were: Leaving the bathroom tap running, leaving the oven on and leaving the main entry door open. These were recorded

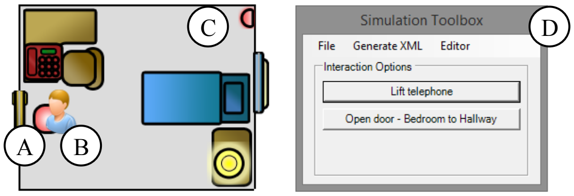

Figure 3. The simulated IE begins generating data once an avatar (B) is placed within it. This example shows a pressure sensor (A) and PIR sensor (C) detecting the presence of the avatar, indicated by a red glow. The simulation toolbox (D) provides a menu facilitating interaction with objects that are within range of the avatar.

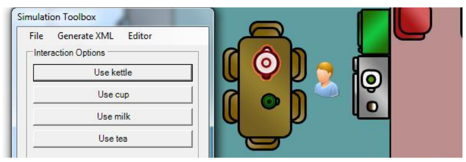

Figure 4. The avatar in the kitchen with a range of the items involved in the "making a cup of tea" ADL.

as individual activities and in parallel with valid ADL performances Fig. 5 provides an example snippet of the homeML data generated during performance of the "make a cup of tea" ADL.

V. Conclusion

Access to data describing the completion of ADLs within IEs is limited due to several complexities associated with the collection of such data. The simulation of datasets provides a method of mitigating these limitations, providing researchers with datasets for the development and testing of novel activity analysis approaches. This paper has demonstrated the use of IE Sim to generate sensor data during the completion of ADLs within simulated IEs using a virtual avatar. This software facilitates the generation of simulated sensor data without the limitations of physical implementations. Environments can be rapidly prototyped with many different environment and sensor layouts prior to physical implementation. Additionally, danger scenarios such as leaving appliances switched on or doors open can be recorded with no danger to participants. The process of simulated IE creation has been illustrated, together with the creation of new sensorized objects. IE Sim's functionality was demonstrated through the recording of the "make a cup of tea" ADL, and its ability to generate diverse and comprehensive datasets in a commonly available format was illustrated through the recording of overlapping activities and potentially hazardous scenarios. Future work on IE Sim will involve the inclusion of support for multiple occupancy simulation in addition to mobile and wireless sensors such as accelerometers. An evaluation will be performed to compare the similarity between simulated and real datasets. There are plans to make the software freely available to the research community via the University of Ulster's Smart Environments Research Group web page [15].


```
...
<location>
   <locationID>0</locationID>
   <locationDescription>Kitchen</locationDescription>
   <locationDevice>
     <ldeviceID>93</ldeviceID>
     <deviceDescription>door</deviceDescription>
     <event>
       <eventID>14</eventID>
       <timeStamp>21:50:20</timeStamp>
       <data>9</data>
     </event>
   </locationDevice>
   <locationDevice>
     <ldeviceID>92</ldeviceID>
     <deviceDescription>pressure sensor</deviceDescription>
     <event>
       <eventID>18</eventID>
       <timeStamp>21:50:24</timeStamp>
       <data>1</data>
     </event>
   </locationDevice>
   <locationDevice>
     <ldeviceID>96</ldeviceID>
     <deviceDescription>cup</deviceDescription>
     <event>
       <eventID>19</eventID>
       <timeStamp>21:50:32</timeStamp>
       <data>9</data>
     </event>
   </locationDevice>
...
```


Figure 5. A snippet of the homeML data generated during the performance of the "make a cup of tea" ADL within IE Sim.

References

[1] United Nations. (2013). *World Population Ageing 2013*. [Online] Available: http://www.un.org/en/development/desa/population/publications/pdf/ageing/WorldPopulationAgeing2013.pdf.

[2] M. E. Pollack, "Intelligent Technology for an Aging Population: The use of AI to Assist Elders with Cognitive Impairment," *AI Mag.*, vol. 26, no. 2, pp. 9-24, 2005.

[3] E. von Strauss et al., "Aging and the Occurrence of Dementia: Findings from a Population-Based Cohort with a Large Sample of Nonagenarians," *Archives of Neurology*, vol. 56, no. 5, pp. 587-592, 1999.

[4] Alzheimer's Society. (2013) *What is dementia?* [Online]. Available: http://www.alzheimers.org.uk/site/scripts/documents_info.php?documentID=106.

[5] E. Tapia et al., "Activity Recognition in the Home using Simple and Ubiquitous Sensors," *Pervasive Computing*, pp. 158-175, 2004.

[6] S. Helal et al., "Persim-Simulator for Human Activities in Pervasive Spaces," in *7th Int. Conf. Intelligent Environments*, Nottingham, UK, 2011, pp. 192-199.

[7] M. P. Poland et al., "Development of a Smart Home Simulator for use as a Heuristic Tool for Management of Sensor Distribution," *Technology and Health Care*, vol. 17, no. 3, pp. 171-182, 2009.

[8] S. Helal, et al., "Scalable Approaches to Activity Recognition Research," in *8th Int. Conf. Pervasive Workshop*, Helsinki, Finland, 2010, pp. 450-453.

[9] A. Mendez-Vazquez et al., "Simulating Events to Generate Synthetic Data for Pervasive Spaces," in *Workshop Developing Shared Home Behavior Datasets Advance HCI Ubiquitous Computing Research*, 2009.

[10] N. Noury and T. Hadidi, "Computer Simulation of the Activity of the Elderly Person Living Independently in a Health Smart Home," *Comput. Methods Programs Biomedicine*, vol. 108 no. 3, pp. 1216-1228, 2012.

[11] A. Helal et al., "3D Modeling and Simulation of Human Activities in Smart Spaces," in *Ubiquitous Intelligence & Computing and 9th Int. Conf. Autonomic Trusted Computing*, Fukuoka, Japan, 2012, pp. 112-119.

[12] C. Krzyska, "Smart House Simulation Tool," M.S. Thesis, Dept. of Informatics and Mathematical Modelling, Technical Univ. of Denmark, Kongens Lyngby, Denmark, 2006.

[13] H. McDonald et al., "The homeML Suite: Shareable Datasets for Smart Home Environments," *Health and Technology*, vol. 3, no. 2, pp. 177-193, 2013.

[14] J. Synnott et al., "IE Sim–A Flexible Tool for the Simulation of Data Generated within Intelligent Environments," in *Int. Joint Conf. Ambient Intelligence*, Pisa, Italy, 2012, pp. 373-378.

[15] University of Ulster. (2014). *Smart Environments Research Group*. [Online]. Available: http://scm.ulster.ac.uk/~scmresearch/SERG/.