Imperial College London

March, 2021

OpenSHS: Open Smart Home Simulator

By Alshammari et al.

Published in MDPI, 2017

Imperial College London

Outline

- Background
- Architecture
- Implementation
- Analysis
- Limitations
- Summary

Imperial College London

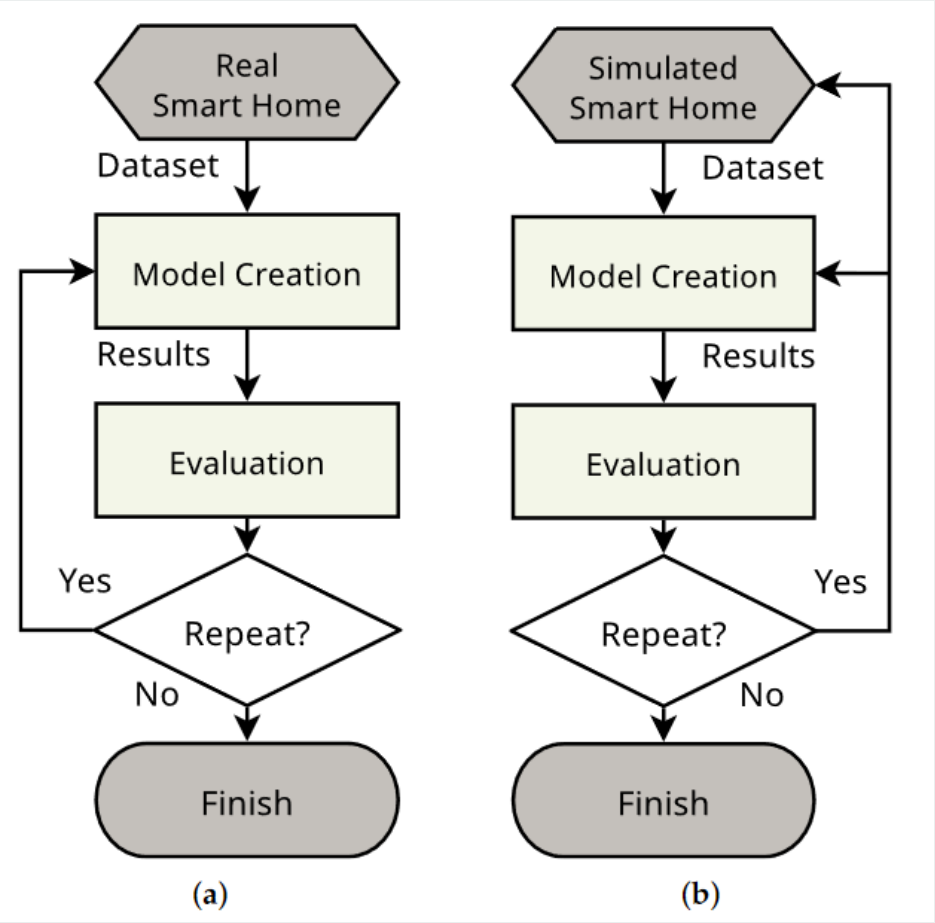
Related work (1)

- Datasets generated by:
  - Real Smart Home Test Beds
    - Need a lot of equipment
    - Takes a lot of time
  - Smart Home Simulation Tools
    - Model-Based approach
    - Interactive approach
    - Most of these models are not available in the public domain

Imperial College London

Related work (2)

- If the results **revealed the need to change** something, this is usually a **costly** and **infeasible** choice to make
  - The researcher can only be able to tweak the model parameters
- With a simulated smart home, this can be easily done, and the researcher can go back and modify the smart home design



Imperial College London

Related work (3)

- Model-Based Approach
  - Uses pre-defined models of activities to generate synthetic data.
    - These models specify the order of events, the probability of their occurrence and the duration of each activity.
  - Facilitates the generation of large datasets in a short period
  - BUT!
    - Sacrifices the granularity of capturing realistic interactions
    - It cannot capture unexpected accidents that are common in real homes

Imperial College London

Related work (4)

- Interactive Approach
  - Can capture more interesting interactions and fine details
  - This approach relies on having an avatar that can be controlled by a researcher, human participant or simulated participant.
  - The avatar moves and interacts with the virtual environment, which has virtual sensors and/or actuators.
  - The interactions could be done passively or actively.
  - BUT!
    - It is a time-consuming approach since all interactions must be captured in real time.

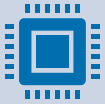
Imperial College London

Motivation

The cost to build real smart homes and the collection of datasets for such scenarios is **expensive**/infeasible

Finding the optimal placement of the sensors, the appropriate participants and privacy and ethical issues

The majority of the existing tools are **not available** in the public domain as an **open-source project**



Most of the publicly-available simulation tools **lack** the flexibility to **add** and **customise** new sensors or devices

Imperial College London

Challenges

- Having a continuous capturing mechanism for the sensors' data.
- An appropriate annotation method for the inhabitants' activities.
- Can they offer:
  - Flexibility and scalability to add new/customised types of smart devices
    - Change their generated output(s), change their positions within the smart home
  - Fast dataset generation
  - The ability to pause and fast-forward the simulation to enable more accurate activity annotation.

Imperial College London

Contribution

- OpenSHS is a new hybrid, open-source, cross-platform 3D smart home simulator for dataset generation
  - Based on Blender and Python
- OpenSHS combines advantages from both interactive and model-based approaches.
  - This approach reduces the time and efforts required to generate simulated smart home datasets
- OpenSHS includes a library of smart devices that facilitates the simulation

Imperial College London

Architecture

Follows a hybrid approach

- Model based approach
- Interactive approach

Three main phases:

- Design
- Simulation
- Aggregation

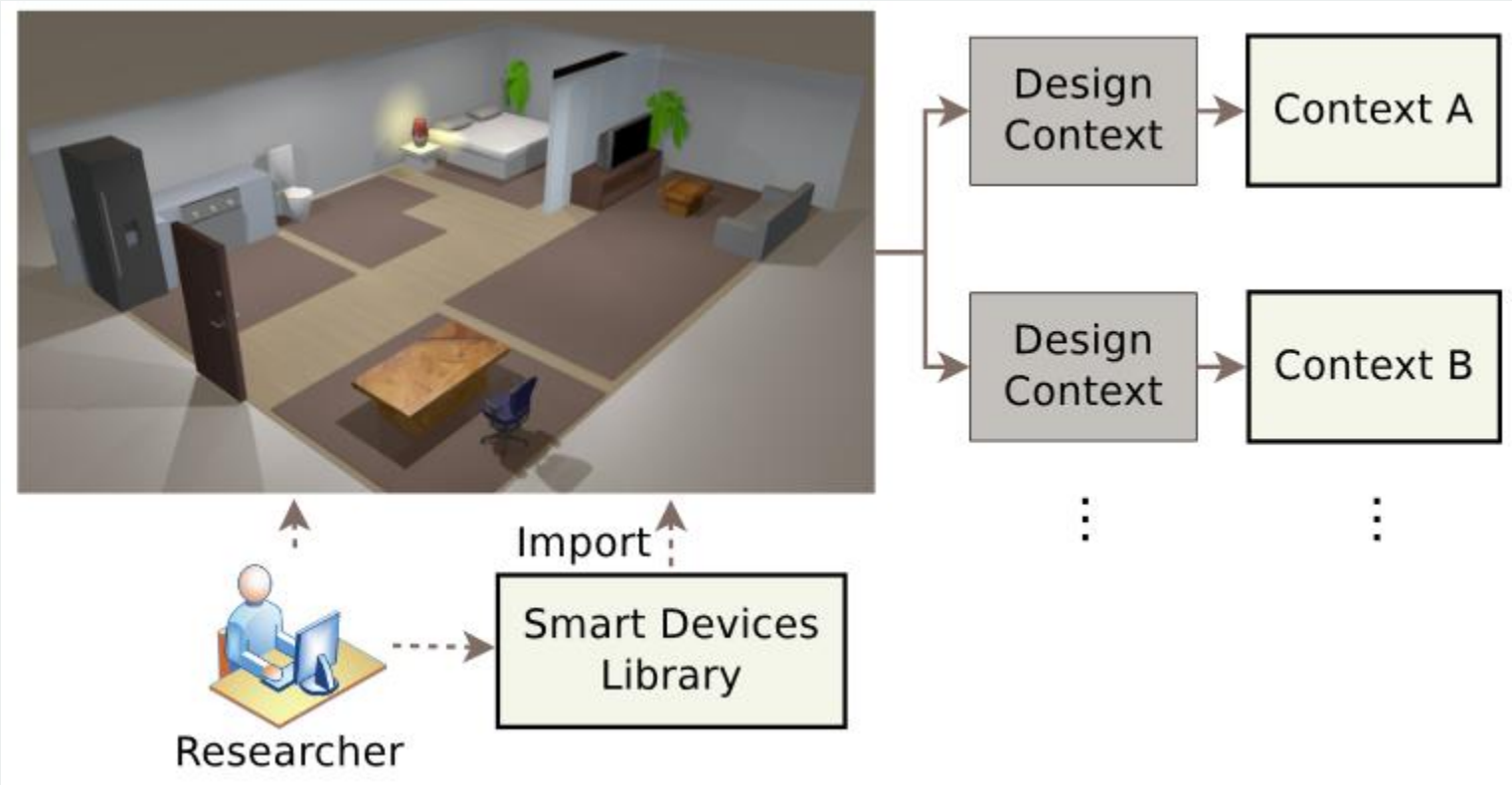
Imperial College London

Experiment

- A virtual smart home environment consisting of:
  - a bedroom
  - a living room
  - a bathroom
  - a kitchen
  - an office
- Each room is equipped with several sensors
  - Totalling 29 sensors of different types.
    - Sample rate: 1 second by default
- The sensors are binary, and they are either on or off at any given time

Imperial College London

Design phase



Imperial College London

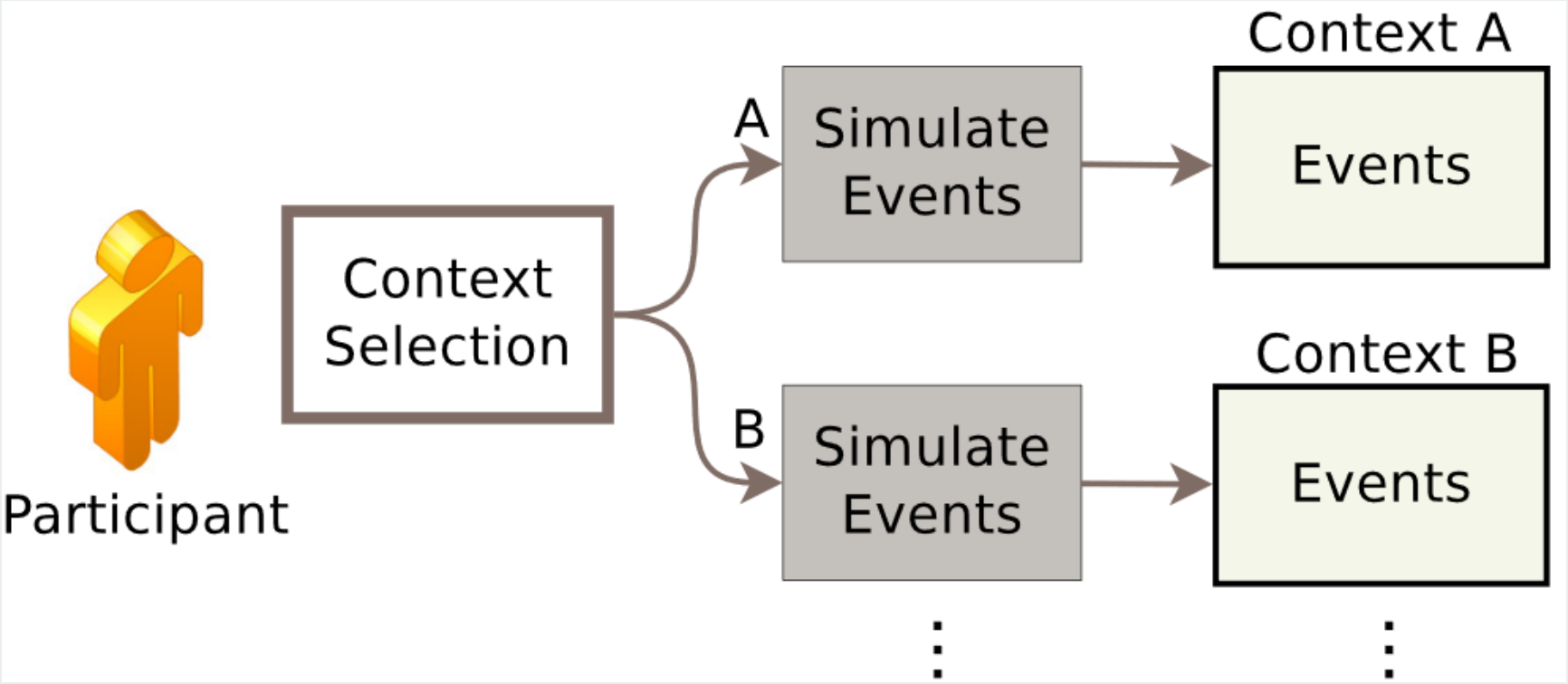
Assigning Activity Labels

- Researchers can define unlimited number of activity labels

- This list represents a sample of activities

Imperial College London

Simulation phase



Imperial College London

Fast forward mode

- OpenSHS allows the participant to control the time span of a certain activity
- The participant wants to watch the TV for a period of time and does not want to perform the whole activity in real time
- The participant can initiate that activity and specify how long this activity lasts.
- The tool will simply copy and repeat the existing state of all sensors and devices during the specified time period.

Imperial College London

Example

- A set of five samples with their activity labels for a certain context.
  - The first sample has five activities and so on
- When the researcher aggregates the final dataset, the samples of every context are grouped by the number of activities in each sample.
  - Sample 1 will be in one group and so on

Samples	Activities				
1	sleep	personal	work	eat	other
2	sleep	personal	other		
3	sleep	personal	other		
4	sleep	eat	personal	other	
5	sleep	eat	personal	other	

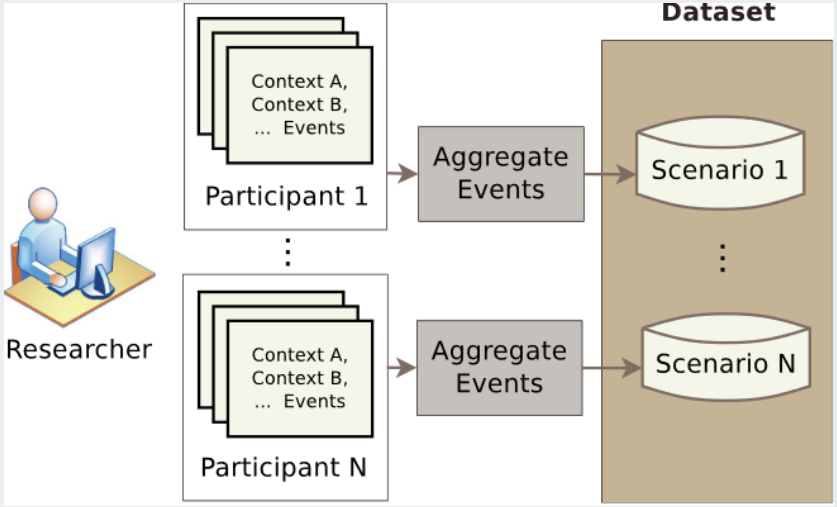
Imperial College London

Activities Labelling

- During a simulation and before transitioning from one activity to another, the participant will choose the new activity from the available list.
- The sample consists of three activity labels
  - ‘sleep’, ‘personal’ and ‘other ’.
- Each activity label corresponds to a set of sensors’ readings.
  - The small circles correspond to an ‘ON-state’ of that sensor.

Imperial College London

Aggregation Phase



- Provides a solution for the generation of large datasets in short simulation time.
- Aggregates the participants' generated sample activities to produce the final dataset.
- The results form a pool of sample activities for each context.
- An algorithm is developed to replicate the output of the simulation phase by drawing appropriate samples for each designated context.
- This feature encapsulates the model-based approach's advantage with the interactive approach adapted by the simulation phase,
  - OpenSHS combines the benefits of both approaches

Imperial College London

Event replication

- Instead of having the user simulating the activities for extended periods of time, the user simulates only a particular context in real time.
- Assume we are interested in an ‘early morning’ context, and we want to capture the activities that the inhabitant is doing in this time frame,
  - What is usually done in the weekdays compared to the weekends in the same context (the ‘early morning’ context).
- The user will only perform sample simulations of different events in real time.
  - The greater the number of samples simulated, the richer the generated dataset will be.

Imperial College London

Replicated copies

- Then a random group will be chosen, and from that group, a sample will be made for each activity.

i	Activity 1	Activity 2	Activity 3	Activity 4	Activity 5
1	**sample 1** sleep	**sample 1** personal	**sample 1** work	**sample 1** eat	**sample 1** other
2	**sample 4** sleep	**sample 5** eat	**sample 5** personal	**sample 4** other	
3	**sample 3** sleep	**sample 3** personal	**sample 2** other		
4	**sample 3** sleep	**sample 3** personal	**sample 2** other		
5	**sample 5** sleep	**sample 4** eat	**sample 5** personal	**sample 5** other	
6	**sample 1** sleep	**sample 1** personal	**sample 1** work	**sample 1** eat	**sample 1** other
7	**sample 2** sleep	**sample 2** personal	**sample 2** other		
8	**sample 5** sleep	**sample 5** eat	**sample 5** personal	**sample 5** other	
9	**sample 4** sleep	**sample 4** eat	**sample 4** personal	**sample 5** other	
10	**sample 2** sleep	**sample 2** personal	**sample 2** other		

Ten replicated copies based on the samples

Imperial College London

Replication Algorithm

- It is not realistic to aggregate the final dataset by trivially duplicating the contexts samples
- There is a need for an algorithm that can replicate the recorded samples to generate a larger dataset
- The number of unique replicated copies for a single context is calculated by:
  - **G** denotes the number of the groups of unique length of activities
  - **Sg** denotes the number of samples for the group **g**;
  - **A** denotes the number of activities within a sample **Sg**. The total number of unique replicated copies **R** is:

$$\mathcal{R} = \sum_{g=1}^{\mathcal{G}} \mathcal{S}_g^{\mathcal{A}}$$

Imperial College London

Dataset generation

- After running the aggregation algorithm, the researcher can combine all of the scenarios, generated by different participants

Timestamp	Bed Table Lamp	Bed	Bathroom Light	Bathroom Door	...	Activity
2016-04-01 08:00:00	0	1	0	0	...	sleep
2016-04-01 08:00:01	0	1	0	0	...	sleep
2016-04-01 08:00:02	0	1	0	0	...	sleep
2016-04-01 08:00:03	0	1	0	0	...	sleep
2016-04-01 08:00:04	1	1	0	0	...	sleep
2016-04-01 08:00:05	1	0	0	0	...	sleep
2016-04-01 08:00:06	1	0	0	1	...	personal
2016-04-01 08:00:07	1	0	0	1	...	personal
2016-04-01 08:00:08	1	0	1	1	...	personal
2016-04-01 08:00:09	1	0	1	1	...	personal
2016-04-01 08:00:10	1	0	1	1	...	personal
⋮	⋮	⋮	⋮	⋮	⋮	

Imperial College London

Future work

- Include multiple inhabitants support in real time
- More sensors
- Add a floor plan editor
- The labelling of activities is performed by the participant during the simulation phase
  - OpenSHS does not perform automatic recognition of these activities
  - Plan to investigate the possibility of adding automatic recognition of the participants' activities

Imperial College London

Summary

- OpenSHS, a novel smart home simulation tool
- Uses the ability of model-based tools to generate large datasets in a reasonable time
- Keeps the fine-grained interactions that are exhibited by interactive tools
- It can generate seeds of events rapidly
  - A replication algorithm that can extend the simulated events to generate multiple unique large datasets.
- OpenSHS offers partial support for multiple inhabitants.

Imperial College London

Thank you!

Questions?