DRSim: A Cyber Physical Simulator for Demand Response Systems

Tri Kurniawan Wijaya*, Dipyaman Banerjee†, Tanuja Ganu†, Dipanjan Chakraborty†, Sourav Battacharya†, Thanasis Papaioannou*, Deva P. Seetharam† and Karl Aberer*

Abstract—Demand Response (DR) is a mechanism in which electricity consumers alter their demand in response to power grid’s supply and economic conditions. DR programs have the potential to improve resource efficiency, sustainability, grid reliability and economic viability by providing tighter alignment between demand and supply. However, implementing DR program is a non-trivial task as it requires good knowledge of electricity consumption preferences, economic models and contextual factors. Developing such knowledge through real world studies can be expensive and be time consuming. As a result, utility companies have been finding it complicated to analyze potential viability and return on investments of DR programs for various ‘what-if’ scenarios.

To address this problem, we present DRSim – a cyber-physical simulator that allows utility companies to study demand side participation aspects of communities with various practical scenarios. DRSim is based on the principles of agent-oriented modeling of users’ behavior and context. It is able to model the emergent behavior of a community based on real data traces that contain partial information about the environment. DRSim is a highly extensible framework to accommodate new data sources, new analytical functionalities and evolve its modeling power. Feasibility experiments show the modeling and analysis potential of DRSim in practical settings.

I. INTRODUCTION

Towards a more sustainable environment, several governments have defined specific goals for reducing total energy consumption (e.g., EU aims to reduce energy consumption by 20% by 2020). With the emergence of technology to support Demand Response (DR), DR programs are known to be effective for reducing energy consumption and peak demand [1], [2]. These DR systems send an "alter demand" signal to electricity customers to alter their consumption at critical times, e.g. during supply-demand imbalance or in response to grid market price fluctuations.

While residential power consumption forms a substantial portion of total energy consumption [3], [4], little is known today on - how much energy is consumed by a particular activity, what human tasks lead to peak usage times, how consumption varies across different customer segments, and which DR programs (e.g. price based [5] or incentive based) are effective for the specific communities. Moreover, the residential DR programs have a major challenge of receiving a good DR participation from a large number of small electricity consumers, where there is much less control and more autonomy [6], [7].The success of a residential DR program largely depends on proper understanding regarding the driving factors of the energy demand for the community and deriving useful conclusions on the demand dynamics. There are various factors that influence residential energy demand, including electricity consumption preferences of different consumer profiles, associated with their life-style preferences and activity-structures, external factors (e.g. weather, seasons, holidays, sporting and cultural events, etc), internal factors (e.g. gender, price sensitivity, income level, age, geographic location presence of guests, etc.) and the DR communication mechanism including the communication channel and timeliness of DR messages. To analyze the potential viability and return on investments of a DR program, it is important to study various ‘what-if’ scenarios involving economic and social interventions discussed earlier and consumers’ response to such interventions.

Hence, research projects by several governments (e.g. [8], [9]) have started collecting fine-grained power consumption data to enable deeper understanding of user consumption behavior, using Advanced Metering Infrastructures (AMIs), additional context sensors and/or user demographics. However, these studies are limited in scale, variability in consumer profiles, availability of all required data sources and coverage over various possible scenarios. Hence, the information available for required community scale analysis is very sparse.

To address the above, recent efforts have focused on creating simulated models of DR for commercial buildings [10], distributed generation [11], [12] and residential sector [13]–[15]. However, these simulators are not designed to capture the collective knowledge available in datasets from real studies that report various sensory observations.

In this paper, we propose DRSim, an organically extensible cyber-physical simulation framework for studying DR programs on a residential community. Our objective is to enable the DR designer to conduct community-scale evaluations of DR programs with best-effort real-world behavior and consumption models of the community. A key novelty in DRSim is that it effectively assimilates physical models of the community coming from sensory observations in real studies, into the simulation process. DRSim models residential customers at a fine-grained level, fuses complex behavior models (e.g. price sensitivity, attitude against sustainability, etc.) with activity models, along with contextual factors like weather, geography etc. It enables DR Designers to find answers to key questions like - determining baselines, root causes of consumption spikes, optimal demand schedules, right choice of incentives, effectiveness of DR programs etc.

Our key contributions are:

• Novel agent-oriented modeling of a simulator so that it can assimilate physical world data from different information sources and infer a community’s consumption behavior and sensitivities.

*Ecole Polytechnique Federale de Lausanne, Switzerland
†IBM Research, India
• Design and implementation of the framework that offers rich analysis of all DR program phases to stakeholders: (1) community-scale data generation and analysis, (2) community-specific DR program design (3) study effectiveness of DR program by observing response from community.

The remainder of this paper is organized as follows: In Section II, we provide an overview of the related work. In Section III, we describe the design challenges for DRSim and present its overall architecture and simulation models. In Section IV, we showcase the effectiveness of our simulator for a specific user community, and finally, in Section V, we conclude our work.

II. RELATED LITERATURE

Prior Simulators: Conducting large-scale trials of DR programs with appropriate implementation is costly. Hence, one direction of work focuses on developing simulations [16], [17] that provide valuable tools to DR designers to conduct cost-effective evaluations of DR programs. For example, Energyplus [17] is a simulation tool for building energy consumption, considering layouts, sunlight exposure, occupancy etc. Zhou et al. [10] model energy consumption in commercial buildings and simulate the effectiveness of price-based DR mechanisms. Densi [11] and Smart Grid Emulator [16] propose DR simulators for distributed generation and smart grids.

Towards simulators for the residential sector, [14] builds a DR simulator that generates energy consumption data by considering demographic types (single senior, regular, couple etc), house type (small, medium, large, etc.) and active hours at home. Effect of DR signals are simulated by grouping appliances into categories and observing the effect of turning off a select set. Another detailed simulation model of the households is proposed in [15]. Appliance energy consumption behavior and interactions are modeled to simulate the energy consumption behavior of a city. DRSim shares the same goal as the above works. However, unlike prior works, we focus on designing DRSim in order to organically incorporate real-world data (smart meter readings, external context data, activity data) that are being generated in different real-world studies. In this context, the key research challenge DRSim addresses is: How do we combine sparse data from different observations to create models that can be used for simulating community’s energy consumption behavior and sensitivities to DR programs?

Agent Architectures and DR: Few works, like us, have adopted agent-oriented design for DR simulation [12], [13], [18]. For example, Burke et al. [13] proposes two agent types: systemic and individual. Systemic agents are used to simulate direct-load-control programs while individual agents autonomously act as the consumer to take appliance usage decisions under constraints. The choice of agent based architecture helps achieving our overall objectives. Firstly, it helps in designing a simulation framework that provides an interplay between various factors like human activities, external conditions, appliance behaviors etc. Secondly, it provides a framework for aggregating and extrapolating data from the sparse real-world information, systematically at different levels – houses, appliances and users. This makes DRSim powerful, extensible and flexible when partial data is available. Finally, we model different user behaviors in a richer way, considering various sensitivities (price, environmental, behavioral). As more data gets collected, these sources will get assimilated in DRSim, leading to better behavior modeling of consumers and hence more realistic results from the tool.

III. DRSIM FUNCTIONAL ARCHITECTURE AND MODELS

DRSim is designed as an organically extensible cyber-physical simulator that can capture physical models of observations from the sparse data of variably instrumented communities in real studies. For the realistic DR simulation, we need two types of data - (1) observatory data from various sensors that provides information regarding energy consumption and the other related driving factors, (2) different behavioral properties of customers in order to simulate responses to DR signals being sent through different communication channels. Given these requirements, DRSim design has the following goals: (i) modularity, so as to facilitate extensibility, (ii) ease-of-use and (iii) support fusion of real and synthetic datasets.

DRSim is designed as a layered architecture (Figure 1).

Data Acquisition Layer: This lowest layer contains various adapters to fuse knowledge from different physical world and synthetic data sources. These data sources are primarily of the following categories: (1) energy data: datasets containing appliance consumption time series data obtained from plug-level meters or house-level meters; (2) activity data: time series of human activity and indoor presence data obtained from variably instrumented homes; (3) surveys containing how customers use appliances; (4) sensitivities: data sources containing information on how users react to demand response signals (e.g., price sensitivity). Synthetic data sources can complement any of the missing models necessary to run DRSim.

Behavior Simulation Layer: This layer models various entities using autonomous and reactive agents. We model three agent types (Figure 2): (i) house agent, (ii) human agent and (iii) appliance agent. Human and appliance agents inherit static properties from their house agent and have their own individual models. This design allows us to easily inject behavior models from observation data and makes the resultant observed behavior per agent and across the community realistic. While human agents are autonomous, appliance agents could be autonomous as well as reactive (i.e. triggered by human agents). Modeling appliances as agents allows us to easily model “smart appliances”, which may have their own

Fig. 1. Functional Architecture of DRSim: The design is inspired by the motivation to feed best-effort models available from physical world observations.
internal behavior model (e.g. for energy savings).

**Simulation Data Store**: The data generated from the simulation runs is fed to applications like baseline prediction, root-cause analysis etc. via the Simulation Data Store. The data consists of a rich, multi-dimensional time-series containing appliance consumption logs, activity logs, aggregate consumption logs of a house-hold, along with static data like demographics. In addition, it contains information about time slots when DR signals were sent.

There are three main technical challenges that we need to address: (a) How do we model the energy consumption behavior of human agents? (b) How do we infer model properties from observation data with varying qualities and sparseness? (c) How do we model a human agent’s response to demand response signals? These challenges are addressed in the following three subsections.

### A. Electricity Demand Model

The first challenge in simulating the community based demand response is to accurately model the electricity demand profiles for different consumers. In real-world scenario, the electricity demand for different consumers can be measured either at household level (using AMI) or at appliance level (using appliance level monitors). We can then build demand models by using the observed data from real-world data sources (by instrumenting certain houses). However, demand models based on such observations alone may suffer from the problem of over-fitting as it completely ignores the underlying human activity which primarily drives electricity consumption in a household. Therefore, it would be prudent to first model human activity which will in turn help us to model the electricity demand by leveraging the activity-appliance relationship. Modeling human activity is a challenging task due to the following reasons:

- The set of possible human activities is very large and varies with time. Additionally, it also differs from one human to another based on human attributes like age, gender, income group etc. Moreover, collection of data related to human activity is costly and tends to be insufficient to represent variations necessary for inductive learning approach to derive any meaningful pattern.

- Inferring appliance usage during an activity is non-trivial as the number and type of appliances used during an activity varies with time and human attribute profiles.

To address these challenges, we include *rooms* as a new variable in our model. We observe that rooms of a house exhibit the following properties:

- The set of rooms are fairly constant for a given house-type and does not change with time. This can be used as domain knowledge.

- Most electrical appliances are stationary and located in a particular room. The appliance-room relationship does not change with time for most appliances and remains fairly similar across households.

- The set of activities performed by individuals have a strong correlation with the room (e.g. cooking in the kitchen). Also, rooms influence appliance usage during an activity from energy consumption point of view (e.g. “reading a book in a porch” and “reading a book in a bed room” are different). This knowledge can be leveraged for better activity prediction. Further, rooms are relatively easier to monitor using sensors compared to activity monitoring.

Figure 3 shows the relationship among different entities in a household. Each household has a fixed set of rooms and different electrical appliances are situated in those rooms. The human-beings in the house perform various activities at different times. At a given time, each activity is performed in a specific room and using a set of appliances situated in that room.

In summary, projection of appliance and activities and their relationship into spatial dimension such as rooms helps to make space of activity and appliances tractable and also improves the predictability of activity-appliance relationship. In addition, spatial states can also help us model user’s behavior w.r.t. DR signals better which we discuss in section III-C. Next we discuss how we use spatial relationship to aggregate and extrapolate the observed data and use it to model demand.

### B. Inference of Model Properties from Sparse Data

We instantiate different agents (the human agents and appliance agents in a given house agent) in the simulation model and model their behavior by estimating probability distributions from real data. We employ a novel supervised machine learning approach for this purpose. Specifically our goal is to infer the following probability distributions:

- \( P(R|H=h) \): used to model the set of rooms \( R \) for a given house type \( h \).

- \( P(D|U=u, H=h, R=r) \): used to model the set of appliances \( D \) present in a given room \( r \) for a human agent profile \( (age, gender, occupation, income etc.) \) \( u \) living in a house \( h \).

- \( P(A|U=u, H=h, R=r, T=t) \): used to model the set of activities a human agent performs given his profile \( u \), in a particular room \( r \) at a specified time \( t \).
and house type $h$.

- $P(D|U = u, R = r, H = h, T = t, A = a)$; used to model the set of appliances $D \in D(r)$ used during an activity $a$ in room $r$ at a given time $t$ by a human agent $u$ living in house $h$ where $D(r)$ denotes the set of appliances already allocated in room $r$ by the simulator using $P(D|U = u, H = h, R = r)$.

As we know, accuracy of any supervised learner depends on the volume of the training instances. Higher the volume, better is the accuracy. However, often it may be difficult to obtain adequate training data when cost of data collection is steep. Hence, we developed a novel incremental imputation technique which is particularly well suited in such cases where the training data is sparse due to high cost of collection. Our technique exploits easily available, time-invariant domain knowledge to impute the sparse data collected from monitoring real users and extrapolates new training instances. We next describe our algorithm formally.

Let us assume there are $n$ attributes which are inter-related and we are given $r$ data sets each representing a relationship among $1 < k \leq n$ of those attributes. We incrementally combine these datasets to infer relations among all the variables that are covered by these $r$ relations and extrapolate new data points using Algorithm 1. It incrementally joins the data sets on matching attributes starting with the those which have the lowest number of matches to achieve generality. The incremental join operations helps to extrapolate new data points by combining observations with domain knowledge while uses the relationships among the attributes to restrict the imputation space and preserve consistency.

**Algorithm 1: Incremental Imputation Technique.**

Input: Set of relational datasets $R = \{R_1, R_2, \cdots, R_r\}$, total number of attributes $n$

Output: Aggregated and Extrapolated dataset $I$

1. $m \leftarrow 2$
2. finish $\leftarrow$ false
3. while finish $= false$ do
   4. $S \leftarrow \emptyset$
   5. $W \leftarrow \emptyset$
   6. for $i = 1$ to $|R| - 1$ do
      7. for $j = i + 1$ to $|R|$ do
         8. if $i < |attrib(R_i) \cap attrib(R_j)| \leq m$ then
            9. $S \leftarrow S \cup \{R_i, R_j\}$ (on all matching attributes)
            10. $W \leftarrow W \cup \{R_i\} \cup \{R_j\}$
   11. $R \leftarrow R - W$
   12. $R \leftarrow R \cup S$
   13. $m \leftarrow m + 1$
   14. if $\forall R_i \in R, |R_i| = n$ then finish $\leftarrow true$

As an example use of our algorithm, let us consider the following datasets collected by monitoring and surveying real users:

- The users’ time-based activity information along with user profile $(U, T, A)$
- Usage of different appliances during specific user activities for their respective user profiles $(U, A, D)$
- House types of different users $(U, H)$

However, as collecting these data is costly, one can use Algorithm 1 to impute these datasets using the following auxiliary relationships which captures how activity and appliances are related with different rooms in a household. This information is comparatively easier to obtain as it is common knowledge and time-invariant.

- User activities “commonly performed at” different room $(A, R)$
- Appliances “commonly located at” different rooms $(D, R)$
- Rooms “found in” different house types $(H, R)$.

The algorithm incrementally joins these datasets to produce the final extrapolated training set consisting of all the attributes $(U, H, R, T, A, D)$ which can then be used to train any supervised classifier and estimate the aforementioned probability distributions required for simulation modeling. Note that, the room attribute $(R)$ is not present in the original observations and is obtained only as part of surveys or domain knowledge. As discussed before, apart from extrapolation, room information also helps to prune the learning space.

### C. Modeling DR Signal Response and Actions

In order to model DR behavior of users, DRSim models:

1. **DR responsivity:** modeling how a human agent reacts on receiving the signal and 2) **DR Action:** modeling agent actions performed while under influence of the signal. Figure 4 provides a representative structure of a DR signal, for example: 

   - **Instruction:** reduce consumption
   - **When:** on weekdays
   - **DR window:** between 5-7pm
   - **Incentive:** get x$/kwh saved

   The message is disseminated in the community using different communication channels at time dissemination time. The DR window is the time window when an agent is requested to take action.

**Modeling DR Responsivity:** We introduce an agent’s DR responsivity $(R_{[0..1]})$ as a parameter that determines whether a human agent intends to participate in DR Action upon receiving a DR Signal. There are several prior arts on modeling agent behavior in response to a stimuli [19]. Given our focus on integration with real-world models and implementation, we define an (extensible) list of principal sensitivity parameters that empirically play a role in decision-making and can be estimated from real-world data or be modeled using distributions. The parameters manifests on both modeling DR responsivity and DR action:

- **energy price sensitivity $R_{[0..1]}$:** This defines how sensitive is the agent against temporal alterations in electricity price. There are many variations of this...
basic parameter (e.g., critical peak pricing, time-of-use pricing) and typical price-sensitivity curves of communities have been reported in prior literature [5].

- **energy perception sensitivity** \( \mathbb{R}[0..1] \): This attribute models how conscious an agent is regarding the adverse effects of energy consumption. It is possible to recover this parameter through surveys and understand metrics that determine green behavior.

- **DR communication sensitivity** A DR Signal is transmitted in several ways – via monthly newsletters, emails; and can consist of real-time messages via smses, in-home displays; or simply community whiteboards. A value of 1 indicates the communication channel is very effective and 0 otherwise. This parameter is dependent on demographic and other attributes of the agent and can be recovered through surveys.

**DRSim implementation** allows definition of new sensitivity parameters and functions, to combine the parameters to determine the final responsivity(t) as a function of time. We calculate the probability of an agent to take DR action during DR Action window by applying a time-based decay function on responsibility(t) between dissemination time and DR window. Note that the DR Action window is different from the dissemination time (see Figure 4). This accommodates users who forget to take action, after demonstrating intention to participate, i.e., if \( \text{Responsivity}_{agent}(\text{DRWindow}) \geq \tau \), where \( \tau \) is a configuration parameter.

**Modeling Actions** Users can perform several actions to reduce consumption; for e.g., turn off un-used appliances, reduce usage time of certain appliances, shift an activity to another time interval. We model two primitive actions in DRSim to implement these actions: (1) \( \text{reduce}(L) \) (2) \( \text{shift}(L) \), where \( L = \text{load} \).

The \( \text{reduce}(..) \) actions are modeled as constraints on two of the three relationships introduced in Section III-A: \( \text{uses}(..) \) and \( \text{performed_in} \). The constraint on \( \text{uses}(..) \) reduces the number of appliances required to perform an activity. This is implemented by dropping appliances that shows low probability of being used in an activity from the observation data. Similarly, the constraint on \( \text{performed_in} \) reduces the number of rooms used for an activity. These two constraints enable us to represent several \( \text{reduce} \) actions. For example, if a regular behavior is to keep lights of few rooms ON while cooking (apart from kitchen), the \( \text{performed_in} \) constraint reduce the number of rooms being used and hence reduce the number of lights being used.

The \( \text{shift} \) actions are used to move an activity from the DR window to another time window. From agent’s activity data, we learn the following constraints on the activity state-transition diagram of a human agent: (1) **quantitative** number of times activity occurs in a time window; (2) **dependency** if \( a_i \) is performed in \( t \), then \( a_j \) is performed as well; (3) **exclusivity** if \( a_i \) is performed in \( t \), then \( a_j \) is not performed. These constraints are used to generate a revised schedule. The specific algorithm used to create a revised schedule is left unspecified in DRSim as many candidates can be used [20]. Finally, we provide an extensible utility loss function to compute user’s inconvenience due to the \( \text{shift} \) operations. This can be used to re-learn user’s responsivity. A simple function to compute user’s inconvenience would be the temporal distance between normal (preferred) and shifted activity schedule along with the relative importance of the appliances that were dropped.

**IV. IMPLEMENTATION AND EXPERIMENTS**

**Implementation Details** DRSim is implemented using discrete time simulation in Java. All physical real-world data sets (discussed further) are loaded into an SQL database. The algorithms in the **Data Acquisition Layer** run in an off-line mode to generate the human activity and appliance behavior models. A simulator kernel maintains rich time context (time, day, month, year) and runs the agents in the user-defined simulation window. We provide an ability to configure various parameters specifying demographic properties of the community being simulated and to analyze the rich logs being stored in the simulation data store. DR events are specified using a comma-separated-values (CSV) format with short-codes to represent different instructions as discussed in Section III.

**Experimental Setup** For our experiments, we focus on simulating a graduate students community, using two real datasets. The first dataset, collected in [21], consists of real activity logs of 5 graduate students, for the duration of approximately one month each. Along with this, we augment a survey dataset by interviewing the same users to obtain their static profiles: gender, age, occupation, and income group, number of rooms and type (kitchen, bed, dining etc), electrical appliances in each room, and their appliance usage pattern, i.e., the set of appliances they use while performing a particular activity. A second dataset provides us with appliance consumption logs from [20], to obtain different operational states of appliances and their corresponding electricity consumptions.

With these datasets, we simulate a community of 100 graduate students households using DRSim. The demographic parameters are configured as follows: (1) **House Type** - 20% studio apartments; 80% single bedroom apartments; (2) **Gender** - 60% males and 40% females; (3) **Family Type** - 10% single, 20% couple and 70% couple with kids.

**A. Modeling Emergent Behavior**

We first demonstrate that DRSim is able to successfully reproduce well-known community consumption behavior. It has been observed before that typical aggregate city-scale residential consumption shows bi-modal distributions. Figure 5 shows the aggregate consumption from our simulated community for two days. We only consider the power consumption due to human activities and omit constantly running appliances such as heater and refrigerator. The first day represents the unperturbed energy consumption behavior of the community and the second day represents the energy consumption behavior as an effect of a DR message. From the Figure 5, it is observed that the aggregate consumption peaks in the morning and evening. This community primarily consists of day-job workers, which also corroborates our result. The fluctuations are due to intermittent use of heavy appliances like stove, oven, vacuum cleaner etc.

Further, we simulate the effect of a DR Signal – **Reduce consumption in the morning (during 9am to 12pm) and receive 1$/kWh saved.** We adopt the incentive sensitivity curve (a log-normal) similar to the one reported in [22] (consumption probability vs. price, Eq. 5) for this purpose, and use our \( \text{reduce} \) models (discussed in Section III-C) to alter the consumption behavior in response to DR. From Figure 5, we observe that there is a drop in the morning peak demand in response to the

---

1[http://www.cres.gr/pepesec/apotelesmata_uk.html](http://www.cres.gr/pepesec/apotelesmata_uk.html)
DR message. Note that these numbers may vary depending on the incentive and community model.

B. Activity and Appliance Energy Consumption

Now we demonstrate DRSim’s ability to simulate realistic activity behavior for human agents and the corresponding energy consumption of various appliances over a day’s time. This simulated data can then offer fine-grained insights to DR designers on root causes of consumption with respect to activities or appliances. Figure 6, shows the probability distribution of different human activities across hours of the day that DRSim learns from the imputed data. Figure 7, on the other hand, shows the relative contributions of different appliances towards the aggregate power consumption at different times of the day.

V. FUTURE WORK AND CONCLUSIONS

In this paper, we described the design and implementation of a DR simulator that can successfully model emergent energy behavior and allow DR designers to simulate demand response signals. Our design is adequately rich to scale to large communities and to accommodate different behavioral models. DRSim model actors as agents, hence it is also possible to model more complex dynamics, such as family with children where appliances are switched on/off erratically. DRSim also allows evaluation of the emergent community-scale effectiveness of DR programs and generates rich data for root-cause analysis of energy consumption. The design goal of fusing real-world knowledge and models leaves a lot of scope for future work. For example, we do not model inter-human interaction and social influence models that can play a role in impacting energy behavior. Also, the impact of a DR stimuli to long term energy behavior of the community needs further investigation. Furthermore, multi-variate root-cause analysis can be complex. As more data gets collected in various projects, we plan to address these issues. Finally, we are going to make DRSim open to the research community for use and further development.

REFERENCES