Virtual Grounding for Facsimile Model Construction
Where Real Data Is not Available

Kotaro Ohori *, Mariko Iida *, and Shingo Takahashi *

Abstract: Recently, a need has arisen for facsimile model construction with close correspondence to the real world for use in evaluating the effectiveness of specific policies. When constructing facsimile models requiring a large amount of data, however, parameter estimation using data fitting becomes problematic. Furthermore, there are often problems where, owing to the limitations of field research, it becomes impossible for analysts to obtain the data necessary for model construction. The paper proposes a solution to this data collection problem by using “virtual grounding” as a method for creating valid agent models. The proposed method constructs an agent model by isolating qualitative features of the real world situation that are targets for modeling at a stage where the complete dataset is not yet available, and uses standard models for which utility has been previously demonstrated. Following this construction, a number of sample participants modeled as agents repeatedly make hypothetical decision-making actions within the model environment, and the model parameters are estimated based on the results of these decisions. This paper demonstrates the utility of the virtual grounding method by using an example of modeling visitor agents to Tokyo Disney Sea.

Key Words: agent-based social simulation, facsimile model, virtual grounding, modeling relation.

1. Introduction

When modeling social systems, grounding of the model with regards to the real world aspects that are the target for modeling allows for determination of model parameters, as well as consideration of the consistency between the behavior of the overall system and real data [1]. Hence, grounding is a part of conventional methods for obtaining validation of a model [2]. Real-world grounding is essential to scenario analysis of specific management conditions, and when performing grounding, actual data serve as the connection between the model and the real world.

Generally speaking, when using simulations as an aid to management, there are a variety of problems concerning the use of data in model validation [3]. In agent-based social simulation (ABSS) research, model validation is confirmed by different methods according to the abstractness of the model (referred to in this paper as the “model resolution”). Validation has a close relationship with the goals to be achieved by applying the model [4]. Gilbert [5] classifies agent-based models into three types according to the validity evaluation: abstract models, middle range models, and facsimile models. Taking the simulation of a cell-type model as an example of an abstract model, an important goal will be to use the simulation results to make meaningful observations related to theory construction for understanding social systems [6]. The validity of the abstract model will therefore be evaluated depending on basic causal correlations that explain the rules of the overall system, and the consistency of these rules with observed events. The middle range model establishes focal points from among complex phenomena, and its primary goal is to analyze the effects of target policies. The validity of the middle range model is evaluated, thereby, according to the simulation results and qualitative consistency between accepted theory and the stylized facts established in the model. Abstract models and middle range models represent not specific social situations, but rather more general ones, and place an emphasis on the discovery of universal social systems. Therefore, even though actual data will be used on occasion for establishing parameters, there are substantial difficulties in evaluating validity through studying the consistency between actual individualized data and the details of agent behavioral models.

The facsimile model, on the other hand, analyzes the utility of specific policies under individualized problem conditions, thus making it an aid in decision-making for management and policy decisions. Identification of agent behavioral models to be included in specific target situations is necessary for effective analysis of specific policies. To that end, actual data must be accommodated in order to calibrate the model and to evaluate its validity.

To date, in the field of social simulation research there have been few reports on policy analysis that uses a facsimile model including detailed and valid behavioral models. The main reason for this lack of studies is that most actual data readily available for use in the creation of agent behavioral models are usually inadequate. For example, to determine the parameters for the Tokyo DisneySea (TDS) visitor agent facsimile model presented in this paper, it was necessary to obtain a sufficient quantity of behavioral data related to visitors under a variety of conditions. Not only do visitors display different behavioral characteristics according to individual differences, but even the same person will show different behavior according to factors associated with the day of the visit (e.g., day of week, weather, and time of arrival). Systematically and exhaustively collecting the amount of behavioral data necessary for agent behavioral model identification is no small task.
To make up for deficiencies in behavioral data, even in cases where analysts perform empirical investigations in an attempt to obtain behavioral model data for agent model construction, limitations on the field research can prevent analysts from collecting sufficient data for facsimile model construction. For example, Tokyo Disney Resorts operate under the concept that they are “the place where dreams come true,” and as a result, the TDS park manager does not permit data collection within the park on the premise that such activities are not a part of visitors’ dreams. This restriction means that detailed data on the behavior of visitors at TDS are unavailable, but also that collecting such data directly from park visitors is not possible. Therefore, demonstrating the validity of specific policies under actual conditions has been difficult, and scenario analysis has been limited to assessing the general effectiveness of policies by using a middle range model.

This paper proposes virtual grounding (VG) as a method for constructing valid facsimile models under conditions where the real-world data essential for behavioral model determination cannot be obtained. The proposed method makes it possible to explicitly describe the effectiveness of policies under specific conditions. As an example application of VG, the construction of an agent model for visitors to TDS is presented. In this example, the overall system behavior and the micro dynamics of individual agents are investigated through simulation results, and the effectiveness of VG is demonstrated through a comparison with actual park conditions.

The main features of VG are that model construction is determined through models for which utility has already been demonstrated, and that data are produced through repeated virtual experiments using a web-based questionnaire on decision-making behavior by agents who are the targets of modeling and yet operate within the model. Hence, sufficient data can be obtained for statistical identification of the model parameters. Moreover, behavioral data are directly related to participant behavior within the model, leading to realistic simulation results.

The remainder of this paper is organized as follows. Section 2 describes the differences between general grounding and the proposed VG method, and gives an overview of the procedures used in the VG method. To demonstrate the effectiveness of VG, Section 3 and subsequent sections present construction of a behavioral model for TDS visitors, and simulations of the behavior of actual visitors. Section 3 formulates a theme park model that supplies the basic TDS park layout, including attractions and routes. Section 4 explains the visitor agent model that was created for the TDS example. Section 5 describes the results of a simulation using the proposed model, and discusses the effectiveness of VG based on the recorded behavior of individual agents.

2. Virtual Grounding Method

2.1 Modeling Relation in General Grounding

Broadly put, grounding makes possible consideration of the relationship between the real world, a model, and data (Fig. 1). The specific functions employed depend on the resolution of the abstract, middle range, or facsimile model used, but in essence a model is constructed through abstraction of the real world, and that model leads to understanding through explanation of real-world phenomena. Grounding of the model with the real world is particularly important when constructing a high-resolution model, such as a facsimile model. Data are normally obtained through real-world sampling. Parameters that determine the construction of the model are estimated from data, and the model generates data that can be interpreted in the real world.

2.2 Virtual Grounding

Facsimile models that analyze policies for specific situations are created in line with actual conditions. As presented in this paper, a theme park such as TDS requires determination of the parameters for tens of thousands of agents composed of the behavioral model.

Determination of such parameters from actual data, however, is complex for the following reasons. Firstly, for statistical estimation methods, enough data can rarely be obtained to guarantee sufficient precision. Secondly, even when a certain amount of actual data has been collected, in many cases the dataset will show only behavior within some limited part of a set of constantly changing conditions. Behavioral models require parameter estimations for agent behavior under a variety of conditions, and a model that generates behavior valid for a given dataset will almost certainly not generate valid behavior under different conditions. Thirdly, real data often cannot be obtained for political or organizational reasons. In the case of TDS, to generate an agent behavioral model for attraction selection, utility function selection parameters are needed to be estimated from actual data (past or present) on park visitors’ selection behavior. Nonetheless, statistically useful data for this purpose do not exist, and park managers refused to permit collection of new data by, for example, asking park visitors to participate in a questionnaire.

In this paper, the problem of obtaining actual data related to behavioral model grounding for a facsimile model is addressed, and the VG method is proposed for generating virtual data to use in behavioral model identification. VG is effective for situations where it is difficult to obtain actual data from systems targeted for model analysis.

Figure 2 shows an overview of the modeling relationships in VG. An existing and accepted behavioral model determines the structure of the behavioral model that is the target of analysis. Real-world data are not used; instead, participants in a questionnaire are used as agent samples and perform virtual behavior within the behavioral model. This virtually generated data (termed “virtual data”) are then used to estimate behavioral model parameters.

The difference between VG and standard grounding is that data to be applied to the model formulation are not acquired di-
Therefore the validation method of a model based on VG is different from standard one from the viewpoint of validating model structure and its parameters. Modelers firstly should specify the model structure from existing accepted model corresponding to “abstraction of real world” in general grounding. Then they should set the model parameters based on the virtual data instead of “real data.” Participants to be modeled as agents “generate” the data by performing virtual decision-making according to the model structure. Thus, if the model has an established method such as maximum likelihood method for estimating its parameters, the model parameters are logically acceptable.

As the research associated with VG, there have been MAS (Multi-Agent Simulation)/RPG (Role Playing Game) methodology [7],[8] and agent-based participatory simulations [9],[10]. The studies also conducted the participatory experiments to gain the agent behavioral data. In particular, agent-based participatory simulations have the advantage to modify a model structure by recoding the behavior. Their purpose, however, is mainly training and education as a support for negotiation. Thus they do not always have to create a facsimile model as long as the participants achieve the purpose with a middle range model. In fact, they could not provide any estimation method for building a facsimile model. VG can be used to identify statistically the parameters of the facsimile behavioral model to be built.

VG is generally performed according to the following three steps. More detailed description how to apply the steps in an actual situation can be found in the subsequent sections.

Step 1: Model selection
The target for analysis determines the qualitative features of the agent, and a model for which utility has previously been demonstrated within the domain (an accepted behavioral model) is used to establish the structure of the agent behavioral model. The model should be selected so that its parameters can be determined, in the following steps, by some calculation method for estimating them.

Step 2: Virtual data generation
Participants to be modeled as agents “generate” data by performing virtual decision-making and behaviors in virtually created situations. It is vital at this step to gather sufficient samples for identification of model parameters for Step 3, and to have participants make repeated decision-making behaviors under changing conditions by dynamically altering parameters. By doing so, data can be collected from similar-type agents under different conditions that are conceivable within the model; thus, the required parameters can be estimated.

To construct a valid behavioral model, carefully selecting the participant sample space is crucial. It is necessary not only to collect a sufficient number of samples, but also to find participants with attributes that represent the distribution found in reality.

Step 3: Determination of model parameters
Data generated in Step 2 are used to parameterize the behavioral model. The validity of the identification method relies on the behavioral model composition. For example, for logit models frequently used in marketing science, segmentation clustering must be performed and parameter estimations are derived by the maximum likelihood method.

3. Applying VG Method to Theme Park Problem
Here, VG is applied to the theme park problem [11], a typical problem for which ABSS is accepted as valid. In the theme park problem, visitors are presented with attraction wait times, and the effects on congestion reduction are analyzed. The utility of such wait time display policies has been previously demonstrated [12],[13]. Previous research used middle range models as behavioral models for parks and agents, but did not take into consideration the specific conditions of individual theme parks. This lack of customization of the problem makes it difficult for theme park managers to verify the effectiveness of policies under conditions particular to the parks that they manage. Here, TDS is taken as a case theme park for application of VG in constructing a park and agent facsimile model.

Section 3.1 discusses formulation of the theme park model that serves as the location where agents act, using basic information about TDS as model data. Section 3.2 discusses the application of VG to construct a visitor agent behavioral model.

3.1 Theme Park Model
The theme park model is composed of four types of structural elements—attractions, events, routes, and entrance gates—each represented as nodes in a network. Links are established between nodes that can be directly traversed by visitor agents.

Figure 3 shows the nodes and links used in the creation of the TDS park model, determined by using a TDS map that was current as of 1 August 2009.

3.2 Attractions
There are 26 attractions at TDS, represented by the set $A = \{a_i | i = 1, 2, \ldots, 26\}$. Each attraction $a_i$ has the following features: capacity, $c_i$; service time, $st_i$; ride interval, $interval_i$; waiting queued population, $waiting_i$; a waiting agent list, $waiting_list_i$; a riding agent list, $riding_list_i$; and an adjacent route $adjacent_route_i$. Here, $c_i$ represents the number of people that $a_i$ can accommodate, $st_i$ represents the duration that agents occupy $a_i$, and $interval_i$ represents the time interval between starts of occupying $a_i$. TDS has attractions where multiple cars operate simultaneously. For such a case, the attraction is modeled with $interval_i$ and $st_i$ being different. In addition, $waiting_list_i$ is a list of those agents waiting at an attraction, and agents arriving at $a_i$ are added to the end of $waiting_list_i$; $riding_list_i$ is a list of those agents currently riding an attraction, and agents in $riding_list_i$ are removed from the list after $st_i$ has elapsed, whereupon they move to the route given by $adjacent_route_i$, and perform the next attraction selection.
“Fast Pass” tickets are issued for popular attractions as a convenience to visitors. The model defines the set of attractions for which Fast Pass tickets are available as FASTPASS\_At = \{a_i | i = 4, 8, 11, 12, 15, 17, 18, 22\}, and issues Fast Pass tickets in a way that allows preferential entry to attractions at specific times. Fast Pass issuance is performed at an attraction’s adjacent\_route_i. In situations where Fast Pass tickets are issued for an attraction, namely, when a_i ∈ FASTPASS\_At, two additional variables are added: the number of riders with a Fast Pass, FASTPASS\_ride\_number_i; and the waiting list of passengers with a Fast Pass, FASTPASS\_waiting\_list_i. The FASTPASS\_ride\_number_i represents the number of agents within FASTPASS\_waiting\_list_i who can preferentially enter the attraction, and those attractions with Fast Pass tickets (a_i ∈ FASTPASS\_At) insert a number of riders equal to FASTPASS\_ride\_number_i into waiting\_list_i, starting from the head of FASTPASS\_waiting\_list_i. Following this, the remaining number of riders, up to the attraction’s capacity, is moved from waiting\_list_i to riding\_list_i.

3.3 Events

Nine parades and shows are represented as a set of events \( E = \{e_k | k = 1, 2, \ldots, 9\} \). A given event, \( e_k \), is performed at a specific predetermined time, and has the following variables: fixed event capacity \( c_k \), service time \( s_k \), an adjacent route adjacent\_route_k, and a list of agents watching the event watching\_list_k.

Agents visiting \( e_k \) are sequentially added to watching\_list_k until \( c_k \) is reached, after which no more agents can watch the event. After \( s_k \) has elapsed from the start of the service, all agents are removed from watching\_list_k and moved onto adjacent\_route_k.

3.4 Routes

Routes represent the surroundings of attractions, events, and entrances. Wait times are displayed and Fast Pass tickets are issued on routes, and this is also the location where agent decision-making takes place. The set of all routes is expressed as \( R = \{r_l | l = 1, 2, \ldots, 27\} \). The time required to travel from the \( l \)th route \( r_l \) to \( a_i \) is at\_time\_\( r_l \) (i = 1, 2, …, 26); similarly, the travel time to \( e_k \) is ev\_time\_\( k \) (k = 1, 2, …, 9) and the travel time to another route \( r_m \) is route\_time\_\( l \) \( m \neq l, m = 1, 2, \ldots, 27 \). Each adjacent\_route_k indicates the wait time \( w_t_h \) (i = 1, 2, …, 26) for \( a_i \), and agents perform decision-making as determined by the wait time reference probability, at\_congestion\_rate_i. The information board variable, Information\_board\_i, indicates the presence or absence of an information board, with Information\_board\_i = 1 indicating that the board is present and, therefore, that an agent can learn the wait time for all attractions. Each \( r_l \) contains the following variables: routes adjacent to the selected route, adjacent\_route_l; adjacent attractions, adjacent\_attr_l; wait time for \( a_i \), \( w_t_h \) (i = 1, 2, …, 26); presence of an information board, Information\_board\_l ∈ \{0, 1\}; and wait time reference probability for \( a_i \), at\_congestion\_rate\_l.

When Fast Pass tickets are available for an attraction adjacent to a route, in other words, when adjacent\_attr_l ∈ FASTPASS\_At, Fast Pass tickets are issued. For such cases, \( r_l \) has the following parameters: number of Fast Pass tickets issued, FASTPASS\_issued\_l; maximum number of Fast Pass tickets that can be issued, FASTPASS\_max\_l; usage time for Fast Pass tickets, FASTPASS\_usage\_l; and number of tickets with updated Fast Pass usage times, FASTPASS\_update\_l. Each time a Fast Pass ticket is issued, FASTPASS\_issued\_l = FASTPASS\_issued\_l + 1. Fast Pass tickets are issued until FASTPASS\_max\_l is reached. Each time a number of tickets equal to FASTPASS\_update\_l is issued, FASTPASS\_available\_l is increased by 5 (FASTPASS\_available\_l = FASTPASS\_available\_l + 5).

3.5 Entrance

Entrance gate \( E_n \) generates agents according to a programmed set of rules, and deletes departing agents. \( E_n \) has only one parameter, an average arrival rate \( \lambda \), determined according to a Poisson distribution. \( \lambda \) represents the average visitor arrival rate for each simulation step.

3.6 Parameters Setting for Theme Park Model

Parameters for the attractions, routes, and entrance gates were set to closely match the actual park conditions. Values were determined through an interview with a representative of Oriental Land Co., the managing company of TDS, and through information made publicly available, for example, on the TDS website. We set parameters values used for service times, visitor capacities, and attraction information such as ride intervals and Fast Pass capacities in the theme park model. Then we identified parameters used for issuing of Fast Pass tickets; and Table 1 shows agent visitor arrival rates and average total agent entries.

4. Visitor Agent Model

This section describes the construction of a visitor agent model according to the three VG steps described in Section 2. Firstly, an attraction selection behavioral model is formulated, based on an accepted model of consumer purchases from the marketing science field (Step 1). Next, the selected model is used to build a model that dynamically varies its parameters according to park conditions, and the participants perform virtual behavior within the model environment (Step 2). Significantly, the method takes into consideration effects of various
park conditions on agent behavior and is one of the most pertinent features of VG as proposed in this paper. To reflect these issues, attributes affecting agent behavioral differences, according to conditions such as day of week (weekday versus weekend or holiday) and accompanying agents (family versus friend), must be accounted for. Therefore, the model parameters are dynamically varied through a web interface, and the participants repeatedly perform decision-making under each condition. The virtual dataset obtained through the repeated decision-making performed in Step 2 is then used to determine visitor agent behavioral model parameters through statistical estimation (Step 3).

Step 1: Model formulation

In Step 1 the visitor agent model is selected by using a model determined to be appropriate through previous research. Visitors usually have multiple preferred attractions upon arrival on a given day, and will select attractions to ride on the basis of their current location within the park and attraction wait times. This behavior closely parallels consumer purchasing behavior, where consumers select from multiple recollected products after narrowing down the selection group and make a selection based on a calculation of product utility [14],[15] (Fig. 4). The present research thus takes advantage of this similarity in visitor and consumer purchasing behavior to create a visitor agent model based on a consumer purchasing behavioral model for which effectiveness has already been demonstrated in marketing science.

Agent’s internal model

A visitor agent, \(v_h\), contains within its internal model both objective and subjective information related to the selection of attractions, and performs decision-making in reference to such information. When each agent arrives at the park, it evokes a list of attractions and events that the agent wants to enjoy, represented as an attractions evoked set \(at_{evoked \_set}\) and an events evoked set \(ev_{evoked \_set}\). Agents choose attractions to actually ride from the evoked attraction list based on wait times and Fast Pass ticket availability. Chosen attractions are represented as \(at_{choice \_set}\). Whether the wait time for \(a_i\) is read from the attraction’s information board is determined according to the parameter \(awareness_{hi}\) \(\in\{0,1\}\) \((i = 1, 2, \ldots, 26)\). Furthermore, each agent contains the following parameter group as an internal model: congestion information, \(congestion_{infohi} \in N\) \((i = 1, 2, \ldots, 26)\) provided via information boards about attraction \(a_i\); a decision-making threshold related to congestion, \(threshold_{hi} \in N\) \((i = 1, 2, \ldots, 26)\); information board reference probability, \(IB_{watch \_rate_h} \in [0,1]\); a set of attractions for which the agent is aware of Fast Pass ticket availability, \(FASTPASS_{aware_h}\); a set of attractions for which Fast Pass tickets are possessed, \(FASTPASS_{possess_h}\); a set of Fast Pass times for each attraction, \(FASTPASS_{time_h} \in N\) \((i = 1, 2, \ldots, 26)\); a list of event starting times for each \(e_k\), \(entertainment_{time_k} \in N\) \((k = 1, 2, \ldots, 9)\); fixed interest level for \(a_i\), \(\alpha_{hi}\) \((i = 1, 2, \ldots, 26)\); and distance and congestion utility weightings for individual agents, \(\beta_{1h}\) and \(\beta_{2h}\), respectively.

Upon creation, agents are assigned \(at_{evoked \_set}\) and \(ev_{evoked \_set}\), in addition to \(threshold_{hi}\), \(congestion_{infohi}\), a fixed \(\alpha_{hi}\) for \(a_i\), and \(\beta_{1h}\) and \(\beta_{2h}\). These parameters are determined in VG Step 3.

Behavioral model

An agent generated at entrance gate \(E_n\) behaves according to the following decision-making process.

(a) Obtaining and updating congestion information

When located at \(r_i\) for which an information board exists \((Information\_board_i = 1)\), \(v_h\) obtains \(IB_{watch \_rate_h}\). Where no information board exists on \(r_i\), congestion information is obtained according to \(at_{congestion \_rate}\) for the adjacent attraction to \(r_i\). At that time, \(awareness_{hi}\) representing knowledge of the congestion information for \(a_i\), is set equal to 1 if congestion information was obtained for that attraction, and 0 otherwise.

Next, updating of \(congestion_{infohi}\) for attractions within \(at_{evoked \_set}\) for which \(awareness_{hi} = 0\) is performed by using congestion information that is presented on other information boards. For cases where there is a mismatch between the time indicated on information boards and the currently held congestion information \((wt_{li} \neq congestion_{infohi})\), \(congestion_{infohi}\) is updated according to

\[
congestion_{infohi} = congestion_{infohi} \times \Lambda_{hi},
\]

where \(\Lambda_{hi} = \frac{\sum_{i \neq hi} \sum_{j} aware_{hi} (\sum_{i} aware_{hi} \neq 0)}{\sum_{i \neq hi} \sum_{j} aware_{hi} (\sum_{i} aware_{hi} \neq 0)}\) \(\times wt_{li} \neq congestion_{infohi}\).

(b) Creation of the selection group

To select attractions that will be actually ridden, \(at_{choice \_set}\) is created by narrowing down \(at_{evoked \_set}\) according to known park conditions.

An attraction \(a_i\) belonging to \(at_{evoked \_set}\) is added to \(at_{choice \_set}\) when the following conditions are met.

1. For each attraction for which the agent possesses a Fast Pass ticket, \(a_j \in FASTPASS_{possess_h}\), evaluation as to whether \(a_i\) can be ridden without exceeding the Fast Pass usage time is performed from

\[
FASTPASS_{time_{hi}} \geq congestion_{infohi} + st_i + at_{time_{hi}} + at_{time_{ij}}.
\]

Here, \(FASTPASS_{time_{hi}}\) is the Fast Pass scheduled time; \(congestion_{infohi}\) is the congestion information for \(a_i\); \(st_i\) is the
ride duration; at \( \text{time}_k \) is the time required to travel from current location, \( l \), to \( a_i \); and at \( \text{time}_{ij} \) is the time required to travel from \( a_i \) to \( a_j \) for which the Fast Pass ticket is possessed.

Agents also evaluate whether there is enough time to view an event after riding an attraction. For each event \( e_k \in \text{ev}_\text{evoked}_\text{set}_b \) (\( k = 1, 2, \ldots, 9 \)):

\[
\text{entertainment \text{time}}_{bk} \geq \text{congestion \text{info}}_{bk} + \text{st}_i + \text{at \text{time}}_k + \text{en \text{time}}_{ik} .
\] 

2. When agents know that \( a_i \) issues Fast Pass tickets (\( a_i \in \text{FASTPASS}_\text{Aware}_b \)) and the agent is not in a time period when Fast Pass tickets cannot be issued, or alternatively, agents do not know that \( a_i \) issues Fast Pass tickets (\( a_i \notin \text{FASTPASS}_\text{Aware}_b \)) or threshold \( \text{hi} \geq \text{congestion \text{info}}_{hi} \).

(c) Attraction selection

Next, the utility value of each attraction in \( \text{at \text{choice \text{set}}}_b \), based on congestion \text{info}_{hi} \) and travel time from \( r_l \) to \( a_i \) is calculated by agents using

\[
U(i) = a_{hi} + \beta_1 \cdot \text{at \text{time}}_b + \beta_2 (1 - \delta_i) \cdot \text{congestion \text{info}}_{hi}.
\]

where \( \delta_i \) = \begin{cases} 1 & \text{if } a_i \in \text{FASTPASS}_\text{At} \\ 0 & \text{otherwise} \end{cases}

The agents’ selection of and movement to attractions are stochastically generated by a multivariate logit model:

\[
p(i) = \frac{\exp U(i)}{\sum \exp U(n)} (a_i \in \text{at \text{choice \text{set}}}_b) .
\]

(d) Riding the attraction

Attraction selections are performed upon each transfer to a route. After arriving, the agent will ride the attraction if \( W_{hi} \) is less than threshold \( hi \).

(e) Receives a Fast Pass

If the wait time is greater than the threshold and the attraction issues Fast Pass tickets (\( a_i \in \text{FASTPASS}_\text{At} \)), then the agent receives a Fast Pass.

(f) Schedule confirmation and random walk

When the attraction choice set is empty (\( \text{at \text{choice \text{set}}}_b = \{ \} \)) the agent does not wish to ride an attraction, and will wander randomly through the park until a scheduled Fast Pass or event time is reached, or until some change in the park conditions is recognized. A route contained within adjacent route, for current route \( r_l \) is randomly selected, and the agent moves to the selected route. After the move, \( \text{threshold}_b = \text{threshold}_b + 1 \) and \( \text{explore \text{time}}_b = \text{explore \text{time}}_b + 1 \). The threshold \( lb \) parameter is reset to its original value each time an attraction is ridden or an event is watched.

Step 2: Virtual data generation

In Step 2, data are generated to identify the model parameter values. As described in Section 2, VG does not determine parameter values from existing data as in traditional methods, but instead generates virtual data according to the behavioral model constructed in Step 1.

To generate virtual data, we chose sample participants that had previously visited TDS, and used an online questionnaire to have the participants perform decision-making behaviors in the constructed model under a variety of dynamically changing conditions. The sample included 1500 participants—enough to produce statistically valid parameter estimates—from an age distribution similar to that of previous TDS visitors. Questionnaire participants were remunerated, as an incentive to provide accurate responses. The following describes the main methods used for determining each model parameter. Figure 5 shows a sample hard copy of the user interface of the online questionnaire.

Evoked set identification

Following the model, each participant selects the events to watch and attractions to ride that most interested them from a list of all TDS attractions. These data were used to determine \( \text{at \text{evoked \text{set}}}_b \) and \( \text{ev \text{evoked \text{set}}}_b \).

Threshold and congestion information identification

For each attraction in the evoked set, participants gave their anticipated wait times and the maximum time that they would be willing to wait in line. The answers were used to establish how the choice sets were narrowed down from the evoked sets.

Utility function identification

By using a map representing TDS attractions and routes, attractions from the evoked set and attraction waiting lists were randomly displayed as shown in Fig. 5. In order that participants would have a feeling of reality, present times, Fast Pass possession information, and Fast Pass usage times were also presented as additional information. From this information, participants virtually performed attraction selection while at the same time viewing actual theme park maps. Each time an attraction was selected, the presented information was dynamically altered. Decision-making behaviors were studied five times for each sample. Data obtained in this manner were then used to resolve the fixed interest level, \( a_{hi} \), for attractions \( a_i \), and the distance and congestion weights, \( \beta_1 \) and \( \beta_2 \), in Eq. (4).

Step 3: Estimation of model parameters

Step 3 estimates each of the parameters comprising the model. First, the 1500 samples, having an age distribution from 18 to 59, were segmented according to the attractions in the evoke lists. Before conducting the main questionnaire survey for VG, we conducted a brief screening questionnaire. First we gained the rates of people having experience of visiting TDS in different ages, which are 18–19, 20–29, 30–39, 40–49, 50–59.
Then we multiply the rate by the number of Japanese population in each age. As a result, we identified the age distribution of the sample participants. To do the segmentation the samples were enumerated using Type III quantification, based on those attractions identified as possible for selection in Step 2.

Second, clustering was performed for those sample scores with cumulative contribution ratios of at least 60%, allowing for classification into three daytime segments and four evening (after 6:00 pm) segments.

Next, by the maximum-likelihood method the attraction selection decision-making data obtained in Step 2 were used to estimate the values of the parameters in Equation (4): \( \alpha_{hi} \) for each attraction \( a_i \), and \( \beta_{1h} \) and \( \beta_{2h} \). The parameters estimated were validated by the significant result of Likelihood-ratio test. The detail result of the parameter values is omitted here due to the page limitation.

Finally, the data obtained in Step 2 were used to establish visitor agent generation rates for each segment, \( \text{threshold}_{ai} \), and the initialization values for \( \text{congestion}_{info} \). The simulation presented in Section 5 is performed by using steps of 1 min, so that steps 1 through 540 represent the segments from 10:00 am to 6:00 pm, and steps 541 and beyond represent the evening segments after 6:00 pm.

5. Theme Park Simulation

Results from running the theme park simulation using the identified parameters are now shown. The example simulation used here is for a typical weekday in August during summer vacation. We verified the effectiveness of the proposed model through comparison between the congestion information in the results of this simulation and those of previous research (Tone and Kohara, 2007) and actual TDS data. Moreover, we investigated the actual micro dynamics of an individual agent behavior.

The simulation was performed with actual TDS data as the basis for visitor agent park entrances.

Figure 6 shows the average wait times for attractions issuing Fast Pass tickets and for other attractions as per a simulation using the proposed model.

Actual data on congestion at TDS for an August weekday during summer vacation indicate that congestion became concentrated at the most popular attractions (those issuing Fast Pass tickets), with other attractions having wait times of approximately 5–20 min. Wait times decreased at times when parade events occur (at steps 180, 420 and 660). This represents a stylized fact that should be shown in our simulation results for confirming the model validation. The simulation results shown in Fig. 6 appear to recreate the actual conditions observed at TDS. Consequently we can conclude that the model constructed in this paper is partly valid also from the viewpoint of model outputs.

Figure 7 shows the peak waiting times for each attraction, comparing the results of the proposed simulation with that of the simulation described by Tone and Kohara [13] and with actual TDS data from 5 August 2011. For comparison with the simulation, we selected TDS data from a date when there were a similar number of park visitors as simulated visitor agents. Some attractions (1, 5, 6, 13, 24, 25 and 26) are not listed, owing to the limited availability of their actual data (in itself an indicator of the necessity for the proposed VG model).

Figure 7 shows that a simulation using the middle range model from previous research indicates extreme waiting times for attractions issuing Fast Pass tickets (attractions 4, 8, 15, and 17), whereas wait times for other attractions are nearly zero—a significant deviation from actual data. In addition, the graph indicates that the simulated waiting time results by the proposed model are close to the actual times for all attractions.

Previous ABSS research has performed model parameter determination after calibration designed to fit macro-scale data, such as market share and sales, or stylized facts. The proposed method, however, uses VG to determine parameters without conducting the calibration. Nonetheless, the model produces simulation results close to actual attraction waiting times, demonstrating the utility of VG.

As has been pointed out recently with regards to ABSS research [4], the utility of a model cannot be fully demonstrated through adjustment of overall simulation behavior to agree with actual data. Model utility should be verified in terms of whether the decision makers could interpret the simulation results as their situations in order to perform scenario analysis with specific policies. Specific scenario analyses are not addressed here because they are beyond the scope of this paper. However, the ability to interpret agent behavior as that of actual park visitors would indicate its value in performing scenario analysis. To verify the interpretability of the behavior of individual agents, therefore, we investigated the micro dynamics of individual agent behavior.

Table 2 shows the behavior of an agent during its day spent in the park.

Table 2 shows that visitor agents select attractions with pref-
of models tuned to those conditions. As an application example, the paper used an accepted behavior selection model from marketing science to construct an agent model for visitors to TDS, extracting samples from an agent sample space composed of 1500 participants. Each participant used an online questionnaire to perform virtual behavior within the selection model under dynamically changing environmental settings, allowing for statistical estimation of behavioral model parameters. A simulation was performed using the behavioral model, and the results were used to examine overall system behavior, as well as the micro dynamics of an individual agent. The utility of VG was demonstrated through comparison of the results with actual park situations.

The park and behavioral models outlined in this paper are specific to TDS, and cannot be directly applied to other situations. Nonetheless, following the method for model construction allows for easy creation of a model for another amusement park with similar features. In that sense, the presented park and behavioral models retain general versatility.

Because the objective of the present paper was to demonstrate the VG method, we have not discussed the analysis of specific policies in depth. However, it is considered a straightforward extension of the model to perform scenario analysis on the effectiveness of policies related to lessening congestion in actual park circumstances. Previously, analysis for cases in which data acquisition is difficult was limited to the use of middle range models; conversely, the VG method as presented here expands the possibility of employing facsimile models that can depict individual circumstances, which would have been difficult in the past.

6. Conclusion

The present paper proposed “virtual grounding” as a grounding method for constructing valid facsimile models where real data for behavioral model parameter identification are not available. This method makes it possible to specifically demonstrate the validity of policies under varying conditions through the use

<table>
<thead>
<tr>
<th>step</th>
<th>Selection behavior (1 step=1 minute)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Enter park.</td>
</tr>
<tr>
<td>14-16</td>
<td>Queue for attraction 19: ride.</td>
</tr>
<tr>
<td>26-34</td>
<td>Queue for attraction 9: ride.</td>
</tr>
<tr>
<td>41-63</td>
<td>Queue for attraction 11: ride.</td>
</tr>
<tr>
<td>138</td>
<td>Obtain Fast Pass for attraction 17.</td>
</tr>
<tr>
<td>148-169</td>
<td>Queue for attraction 3: ride.</td>
</tr>
<tr>
<td>170</td>
<td>Wander around park.</td>
</tr>
<tr>
<td>192</td>
<td>Queue for attraction 10: ride.</td>
</tr>
<tr>
<td>241-281</td>
<td>Queue for attraction 22: ride.</td>
</tr>
<tr>
<td>297-361</td>
<td>Queue for attraction 6: ride.</td>
</tr>
<tr>
<td>377-395</td>
<td>Queue for attraction 7: ride.</td>
</tr>
<tr>
<td>396</td>
<td>Wander around park.</td>
</tr>
<tr>
<td>458</td>
<td>Queue for attraction 2: ride.</td>
</tr>
<tr>
<td>540-599</td>
<td>Queue for attraction 13: ride.</td>
</tr>
<tr>
<td>607-654</td>
<td>Queue for attraction 5: ride.</td>
</tr>
<tr>
<td>655</td>
<td>Wander around park.</td>
</tr>
<tr>
<td>704</td>
<td>Queue for attraction 17: ride using Fast Pass.</td>
</tr>
<tr>
<td>716</td>
<td>Leave park.</td>
</tr>
</tbody>
</table>

References


Kotaro OHORI (Member)
He received his Ph.D. from Waseda University in 2011. He was an assistant professor at Waseda University till 2011, and is currently a researcher at Fujitsu laboratories Ltd. His research interests include evolutionary economics, service science, complex networks science, and agent-based modeling and simulation.

Mariko IIDA
She received her M.Eng. from Waseda University in 2011. She is currently working at Media Department, Leverages Inc.

Shingo TAKAHASHI (Member)
He received his Ph.D. from Tokyo Institute of Technology in 1989. He is currently a professor at Waseda University. His research interests are mainly in social systems science including agent-based social simulation and soft systems approach.