VALIDATION OF AGENT-BASED MODELING: CASE STUDY FOR A SYNTHETIC LABOR MARKET

Alok Chaturvedi, Purdue University, West Lafayette, IN 47907, alok@purdue.edu
Daniel Dolk, Naval Postgraduate School, Monterey, CA 93943, drdolk@nps.edu
Donald Kridel, University of Missouri-St. Louis, St. Louis, MO 63121, kridel@umsl.edu

INTRODUCTION

Agent-based modeling (ABM) has enjoyed rapid growth in the past decade, especially with respect to the development of virtual environments such as computational organizations [1], synthetic economies [5], and social networks [7] [2]. One of the challenges with agent-based systems is model validation. Since these agent-based environments often consider scenarios of "what might be" or "what should be" in addition to "what is", and since the system phenomena "emerge" from numerous nonlinear interactions of agents with their environment and with each other, it is sometimes difficult to determine the validity of the underlying ABM. We use the methodology of docking to compare an ABM with a discrete choice model to see whether they yield similar results.

VALIDATION OF SIMULATION AND AGENT-BASED MODELS

Validation for simulation models has been implemented using a variety of methods such as:

- ✔ Evaluation of the model by subject matter experts;
- ✔ Use of a Turing test;
- ✔ Use of confidence intervals and hypothesis tests;
- ✔ Comparison of model output data with comparable data from a similar existing system;
- ✔ Comparison of model output data with the comparable output data from another model that is though to be valid. [6]

The last method of aligning two models is closely related to docking: "Docking juxtaposes two models to investigate whether they proceed in like manner or yield similar results." [3]. We adopt the docking validation approach by comparing a discrete choice logistic model with agent-based simulation results for a synthetic labor economy.

APPLICATION: SYNTHETIC LABOR MARKET FOR ARMY RECRUITING

The ABM we consider for validation is an artificial labor market (ALM) constructed for the U.S. Army Recruiting Command to conduct strategic war gaming and decision support exercises. We present a very general description here; a more detailed description appears in [4].
There is only one class of artificial agent in the ALM, and that is the potential recruit who has already contacted the Army through a recruiter or via the Web, and expressed interest in enlisting. The agent has the attributes (gender, age, education, etc.), or genes, and undergoes the decision process:

- ✔ agent may respond to contact;
- ✔ once responding to contact the agent may sign a contract;
- ✔ once contracted the agent mat may report for active duty or "opt-out" (DEP loss)

The current simulation has one artificial agent for every recruit, or roughly about 1.5 million agents.

The specific stage of the decision process for which we designed and built validation models is the DEP (Delayed Entry Program) decision point. The DEP is the equivalent of inventory in a supply chain model. When an individual signs a contract to enlist in the Army, s/he usually does not go immediately to basic training for a variety of reasons, e.g., the individual may not have graduated from high school yet, there may not be room in the training facilities, etc, etc. Most people who sign a contract go into the DEP until they are ready to be assigned; this may be a period ranging from a week to a year. DEP Loss refers to those enlistees who drop out of the DEP, most often because they have changed their minds about joining, but possibly for other reasons such as encountering trouble with the law, or failing to pass a follow-up physical.

DEP Loss is a critical parameter for Army recruiting because the higher the loss rate, the more work has to be done on the front end by recruiters getting new enlistees through the door. Thus, it is of interest to try to identify a priori the attributes of individuals who are likely to contribute to the DEP Loss, in order to avoid, when possible, investing significant effort in their recruitment.

VALIDATION METHODOLOGY: DISCRETE CHOICE MODELS

ABM decision rules and calibration methodology for DEP Loss

The ALM decision rules for agents are described in [4]. The parameters in these rules have been calibrated to align with historical data from years 1999 through 2004. The aggregation level of alignment is quite high, however. Recall that there is a large number of segments; if we take the combinatorial product of all possible attribute levels, there is a universe of ~45,000 different segments or cells. Through the simulation interface, the users can view not only the aggregate results (total Army, total brigade, total battalion), but can also drill down to any one of these segments and see the results. This is a powerful feature of the simulation but it raises the challenging question of how valid the results are for any particular cell.

Discrete choice models for DEP Loss

Using a database of over 80,000 recruits, discrete choice models have been developed. Appendix 1 provides DEP loss model results.[*] The model can be used to predict the probability that a given recruit would become a DEP loss. Likewise, the individual probabilities may be aggregated to generate predictions for any subset of interest that is of reasonable size. Model results for these subsets of interest can then be compared to the ABM results.

Experimental design

Armed with the results of the DEP Loss logistic models, the challenge now becomes how to compare the ABM results with these models. We outline the following methodology for how to achieve this.

The first realization is that of the 45,000 cells, only a relatively small subset has enough individuals for a comparison to be statistically significant. We therefore identified the population of cells containing 30 or more individuals.

[*] A three-step modeling process that follows the decision criterion has been proposed; funding is not yet available.

The next step is to select a sample of this population of statistically significant cells for comparison. While there are many possible criteria for selection, we used two methods for selection. One, subsets of particular interest to the Army were selected; second, a "broad" cross-section of alternative sub-sets were selected. This methodology provided a way to identify "high-cost" failures and "general" failures in the ABM simulation results.

The next step is to run the ABM a predetermined number of iterations in order to obtain a meaningful distribution of DEP Loss values for each of the sample cells. We set the exogenous economic and force strength requirements parameters to match as closely as possible with the actual values for the year 2004. These simulations will be run 250 times to create distributions for each of the sample cells identified in the previous step.

For each simulation cell distribution, the corresponding logistic model results for that cell within the distribution will be compared to the ABM results; this will allow us to see how far model results deviated from the mean. Situations with large deviates will be "flagged" for further analysis.
Once these ABM/model "mismatches" have been identified, these "mismatches" can in principle be used to alter the utility (decision) functions for the agents. This back propagation technique is an area for further research.

CONCLUSIONS

We have presented an initial methodology for validating an ABM for an artificial labor market using a discrete choice logistic model as its docking complement.

REFERENCES

[1] Anderson, Philip. Complexity theory and organization science. *Organization Science*, 10 (1999), 216-232.

[2] Barabasi, A-L. Linked*: How Everything is Connected to Everything Else and What It Means*. Plume Press (2003).

[3] Burton, R. Computational laboratories for organization science: Questions, validity and docking. *Computational and Mathematical Organization Theory*, 9, 2 (Jul. 2003), 91-108, http://www.kluweronline.com/issn/1381-298X/ .

[4] Chaturvedi, A., Dolk, D., Mehta, S., Ayer, R. Agent-based simulation for computational experimentation: Developing an artificial labor market. *European Journal of Operational Research*, 166, 3 (Nov. 2005), 694-716.

[5] Epstein, J. and Axtell, R. *Growing Artificial Societies: Social Science from the Bottom Up*. The Brookings Institution and the MIT Press, Washington, DC and Cambridge, MA (1996).

[6] Law, A. and Kelton, W.D. *Simulation Modeling and Analysis (Third Edition)*, McGraw-Hill, 2000.

[7] Wasserman, S., Faust, K., Iacobucci, D., Granovetter, M. *Social Network Analysis: Methods and Applications*. Cambridge University Press, Cambridge, UK, 1994.

APPENDIX A: DEP-LOSS LOGIT MODEL

Dependent Variable: DEP_VAR
Method: ML -Binary Logit (Quadratic hill climbing)
Sample: 1 85607
Included observations: 85490

Variable	Coefficient	Std Err	t-statistic	Prob.
C	-0.63978	0.24385	-2.62363	0.01
MALE	-0.42356	0.02882	-14.69865	0.00
EDCLASS1	-0.35857	0.14330	-2.50225	0.01
EDCLASS2	-0.37808	0.14592	-2.59102	0.01
EDCLASS3	-0.33468	0.14612	-2.29044	0.02
EDCLASS7	-0.11293	0.14447	-0.78170	0.43
M_RACEA	-0.14851	0.08171	-1.81762	0.07
M_RACEB	0.09125	0.03450	2.64495	0.01
M_RACEO	0.19596	0.03522	5.56398	0.00
M_RACEP	-0.36489	0.12791	-2.85264	0.00
AGE	0.03276	0.00338	9.70396	0.00
TSC2A	0.37206	0.09712	3.83110	0.00
TSC2B	0.18846	0.09126	2.06513	0.04
ADVRNKC1	-0.85775	0.07008	-12.24012	0.00
ADVRNKC2	-0.48067	0.04710	-10.20467	0.00
ADVRNKC3	-0.81622	0.05973	-13.66465	0.00
ADVRNKC4	-0.49305	0.03975	-12.40514	0.00
ADVRNKC5	-3.12392	0.14120	-22.12355	0.00
RELIGION_CHRISTIAN	-0.21954	0.09128	-2.40507	0.02
RELIGION_CATHOLIC	-0.23226	0.09435	-2.46174	0.01
RELIGION_ADVENTISTS	-0.30404	0.24099	-1.26164	0.21
RELIGION_MORMON	-0.05757	0.14237	-0.40435	0.69
RELIGION_NO_PREF	-0.12868	0.09048	-1.42216	0.16
MM	-0.00766	0.00635	-1.20534	0.23
MVB	0.00098	0.00040	2.42768	0.02
OFSCORE	0.01724	0.01217	1.41641	0.16
GM	-0.02240	0.00743	-3.01673	0.00
ST	0.00109	0.00804	0.13532	0.89
BRIG_3	-0.15589	0.03371	-4.62381	0.00
BRIG_4	-0.26470	0.03404	-7.77665	0.00
BRIG_5	-0.24037	0.03821	-6.29112	0.00
BRIG_6	-0.19260	0.03358	-5.73532	0.00

Mean dependent var	0.112715	S.D. dependent var	0.316246
S.E. of regression	0.311945	Akaike info criterion	0.672637
Sum squared resid	8315.897	Schwarz criterion	0.676139
Log likelihood	-28719.87	Hannan-Quinn crit.	0.673707
Restr. log likelihood	-30105.66	Avg. log likelihood	-0.33594
LR statistic (31 df)	2771.588	McFadden R-SQ	0.046031
Probability(LR stat)	0		
Obs with Dep=0	75854	Total obs	85490
Obs with Dep=1	9636		