Proceedings of the 1985 Winter Simulation Conference
D. Gantz, G. Blais, S. Solomon (eds.)

Optimization of Manufacturing System Simulations Using Perturbation Analysis and SENSE

Rajan Suri
University of Wisconsin
Dept of Industrial Engrg
Madison, WI 53706

Gregory W. Diehl
Network Dynamics Inc.
1218 Mass Ave
Cambridge, MA 02138

Yu-Chi Ho
Harvard University
Divn of Applied Sciences
Cambridge, MA 02138

ABSTRACT

Perturbation Analysis (P/A) of discrete event systems is a recent technique which enables sensitivity analysis of decision parameters from a single simulation run. Thus, from a single simulation of a manufacturing design, one can estimate the effects of changing many different decisions such as adding more machines or fixtures, changing buffer sizes or changing lot sizes. The first part of this paper gives an overview of this new technique. The second part describes SENSE (SENSitivity Evaluator), a software package implementing the theories of P/A in easily usable form. Application of SENSE to GPSS/H simulations is described, with examples of its use for the design of flow lines at a major U.S. corporation. By using SENSE a simulation team is able to improve their designs in less time, and with less manpower and CPU-time, than previously required. The result is better designs and better turnaround time for simulation activities in general.

1. BACKGROUND

1.1 Manufacturing Systems and Modelling

The need for increased productivity in manufacturing is leading to the introduction of sophisticated integrated manufacturing systems throughout industry [1]. However, the complexity of these systems makes them difficult to design and operate efficiently [2]. In the design and operation of manufacturing systems, it is useful to have methods that can predict the performance of the system due to changes in various decision parameters [2,3].

In the early stages of system design and planning, analytic models are useful in deciding the overall system parameter values. An example of such a model being used by designers today is MANUPLAN [4], which is an efficient and easy to use performance evaluation tool at the planning stages. However, in order to proceed to detailed system design and fine tuning of the parameter values, it is necessary to construct detailed simulation models of the systems.

Traditionally, designers of manufacturing systems spend a good deal of time experimenting with these simulation models, to find the tradeoffs between various decision parameters, and generally to gain insight into the performance of the system. During the process of experimentation, if the sensitivities of performance to the various parameters could be estimated, then these tradeoffs would become clear, and the sensitivity values would give a lot of insight into the leverage points of the system, as well as the insensitive parameters.

Conventionally if such sensitivities were desired, they would be obtained as follows. By running the simulation, a performance measure such as production rate is obtained, and by altering one of the decision parameters and repeating the experiment, an estimate of the sensitivity (gradient) of the performance measure with respect to the decision parameter is obtained. This sensitivity estimate can then be used to decide how to improve the parameter value. However, an accurate simulation can be expensive to run, whenever a large number of "what-if" scenarios are to be tried with many parameters. Also, a great deal of the designer's time may be spent both in deciding which combinations of parameters should be varied, and in reviewing the results of all those experiments. In short this process can be time-consuming for the designer as well as the programmer.

1.2 Perturbation Analysis

Recently a set of techniques, known as Perturbation Analysis (P/A) of Discrete Event Systems, has been developed that overcomes the disadvantages mentioned above. When used with simulation, it enables a large number of parameter sensitivities to be estimated from just a single simulation run. P/A can also be used on-line for an actual manufacturing system such as an FMS, to determine parameter sensitivities for the actual system while it is operating -- this point will not be covered in the current paper (see [5] for details). Also, a technical description of P/A will not be given here: instead, we give an overview of the developments in P/A, along with a bibliography of technical papers containing details of the method.

The P/A approach was originally developed to solve a design problem in production lines [6]. Since then, various theoretical and practical extensions have made it applicable to a wide variety of parameters and systems (see literature survey in next section). P/A works by observing a single experiment on the system, called the nominal experiment. This experiment can be obtained from a simulation or by observation of an actual system. While the experiment is evolving, P/A considers what would have happened if various parameters were different, for example, if a machine were faster, or if there were more operators or transport devices. P/A performs simple calculations to obtain answers to these questions in a very efficient manner. The additional calculations require relatively little additional memory or CPU-time in a simulation and are simple enough to be performed on-line as an existing system is operating, or to be performed by subroutines during a simulation run. These calculations provide the desired parameter sensitivity estimates at the end of the observation on the single experiment.

2. BRIEF REVIEW OF PERTURBATION ANALYSIS LITERATURE

This section gives the reader a guide to publications on P/A.

A thorough and quite up to date survey of P/A is provided by Ho et al. [7], and we would recommend that article, with its comprehensive bibliography as a good overview of P/A. Here we will briefly cover the main developments. Much of the early work on P/A concentrated on queueing network models. The basic algorithm for deriving throughput sensitivity with respect to service time parameters in a queueing network is given in Ho and Cao [8]. Generation of perturbations for general models is described in Suri [9]. The algorithm for sensitivity with respect to number of fixtures (equivalent to number of customers in a closed network) is given in Suri and Cao [10]. It is interesting to note that a recent paper [15] showed an independent discovery of the basic P/A approach for closed queueing networks, by a researcher who was apparently unaware of any of the other references cited here.

Next, there are extensions of P/A to other parameters and performance measures. Suri [11] gives an algorithm for general systems and general performance measures, although for a restricted class of parameters. Ho et al. state a technique for buffer size sensitivity [6] as well as other parameters [12] for a production line. Ho and Cao [13] show the application to sojourn time sensitivity estimation. Related to all these is the result by Zazanis and Suri [14] showing that P/A is much more efficient than alternative methods for deriving sensitivity estimates. Finally, there are papers showing the use of the sensitivity estimates derived by P/A, in optimizing parameters for a system. Ho and Cao [8] combine P/A with stochastic approximation methods to optimize parameters for a queueing network. Suri and Zazanis [16] show how P/A can be used to optimize parameters for single-server queues, during a single simulation run.

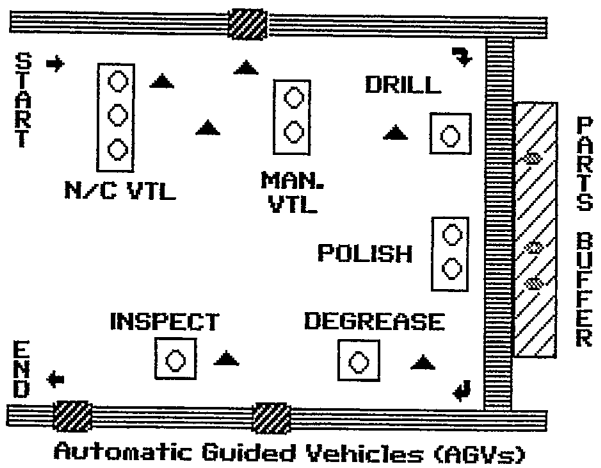
3. APPLICATION OF SENSE TO THE SIMULATION OF MANUFACTURING SYSTEMS

While we have explained above the principle behind P/A as a "simulation enhancer", the actual implementation of the idea in a real world simulation environment is embodied in a software package called SENSE (SENSitivity Evaluator).

To show how SENSE is used we first describe a typical flow line manufacturing system and show a portion of a GPSS/H simulation model of it. We then illustrate how SENSE is incorporated into the model.

3.1 DESCRIPTION OF A MODEL OF A MANUFACTURING SYSTEM

To illustrate the application of SENSE to a particular simulation language (GPSS/H), let us consider a typical flow line manufacturing system (FLMS) shown in Figure 1.



The FLMS consists of:

(i) 1 Machining Station with 3 Numerically Controlled Vertical Turret Lathes (NC/VTL)

(ii) 1 Machining Station with 2 manually Controlled VTLs

(iii) 3 operators shared among the 3 NC/VTLs and the 2 MAN/VTLs

(iv) 1 Drilling Machine Station

(v) 2 Polishing Stations

(vi) 1 Degreasing Tank

(vii) 3 operators shared among the polishers, the degreaser, and the inspection station

(viii) 1 Inspection Station

(ix) 3 Automatic Guided Vehicles (AGV)

(x) 1 Parts Buffer

There are four types of parts to be produced by this FLMS. They enter the system at the first station and are transported to the appropriate machining stations by the AGVs in sequence. If a particular machine is not available for the part in transit, then the part is deposited in the parts buffer temporarily. Parts are produced in lots and do not need to move in a group, but can move individually. In this way a number of machines may be working on the same batch of parts simultaneously. In addition the machines and operations are designed so that parts move from the front of the area to the end in a "continuous flow". The flow aspect and the fact the pieces move independently distinguish the flow line from the JOB SHOP.

Operators are required at various machines for setup, gauging, etc. during the machining operation but not for the cutting portion of the operation. Thus, they can be shared among machines. Operators can be preempted by jobs with higher priority. The performance measures of the system are Work-In-Process (WIP) and Flow or System Time (FT). The system parameters for which sensitivities are sought are: setup times and machining times at various machines; the arrival rates and lot sizes of the different part types; and the number of AGVs.

To illustrate how SENSE is incorporated into a model let us consider a typical segment of the GPSS/H code for the above FLMS simulation. Several part types must all go through a common manufacturing cell involving one machining operation. Each part type must

(i) queue and seize an operator of an appropriate type
(ii) queue and seize a machine of a special type
(iii) have the operator commence machining operation on the part
(iv) release the operator for other duties
(v) upon completion preempt (according to priority, queue if necessary) the operator for a specific length of time
(v) release both the machine and operator and leave the cell

The pertinent section of a GPSS/H program implementing the above, stripped of all input/output and other housekeeping instructions, looks something like the code in figure 2.

```
             GPSS/H INSTRUCTIONS
                  (REMARKS)
                     .
                     .

ENTER      OPER2          get operator.
SEIZE      VTL1           get first machine.
ASSIGN     2,PH2,EXP1,PF
                          specify start time
ADVANCE    PF2            advance start time
                            (with operator)
LEAVE      OPER2          return operator


ASSIGN     2,PH3,EXP1,PF
                      specify machining time
ADVANCE    PF2        advance machining time


PRIORITY   1              set priority, get
ENTER      OPER2          next free operator
PRIORITY   0              reset priority
ASSIGN     2,PH4,EXP1,PF
                          specify end time
ADVANCE    PF2            advance end time

LEAVE      OPER2          return operator
RELEASE    VTL1           release machine
    .
    .
    .
```

FIGURE 2

3.2 APPLICATION OF SENSE TO GPSS/H MODELS

Now imagine that full GPSS/H simulation program for the above FLMS has been coded, and that we wish to install SENSE to evaluate sensitivities.

Recall that perturbation analysis requires that each transaction in the simulation be analyzed and that certain gradient data be derived for the transaction. Thus in a GPSS/H environment each command which involves the initiation or termination of a transaction requires a CALL to the SENSE subroutine. This is accomplished in practice in two ways:

A) People writing the GPSS/H code are instructed in the use of SENSE. As they write code they include a SENSE call before or after the appropriate transaction. Our experience proved this to be somewhat time consuming and distracting. We therefore developed a second alternative.

B) GPSS/H code is written without any consideration of SENSE. When the GPSS/H code is in place and debugged, SENSE is installed in the code using a software instrumenter. The instrumenter places 95% of all the necessary SENSE calls into the GPSS/H code. The person writing the code then reviews the code and instruments the final 5% of the SENSE calls by hand. This has proven to be an efficient way to install SENSE.

To add SENSE capability to a portion of the code, we add various "calls" to the above program segment (figure 2) with the result that the program now looks like the code in Figure 3. The bcall statements are blocks executed by jobs as they move through the system. They invoke the named subroutine and pass the arguments listed to the subroutine as parameters.

Note that all added bcall instructions are starred: these are the calls to the SENSE routines. The parameters associated with the bcalls provide data, such as: job id (previously stored in pf16), clock time (AC1), device name (OPER2, VTL1...) and types of operations being executed (1,2,3.. in the advance calls). Particular GPSS/H instructions (e.g. SEIZE, ENTER ...) require a pair of calls because SENSE needs to know both the time the job first began queueing and when it actually gained control of the resource.

```
                GPSS/H INSTRUCTIONS (with SENSE)
                      (REMARKS)
                  .
                  .

*     bcall      &benter(pf16,ac1,OPER2)
      ENTER      OPER2       get operator.
*     bcall      &aenter(pf16,ac1,OPER2)

*     bcall      &bseize(pf16,ac1,VTL1)
      SEIZE      VTL1      get first machine.
*     bcall      &aseize(pf16,ac1,VTL1)

      ASSIGN     2,PH2,EXP1,PF
                           specify start time
      ADVANCE    PF2       advance start time
                              (with operator)
*     bcall      &aadvanc(pf16,pf2,1)

      LEAVE      OPER2       return operator
*     bcall      &aleave(pf16,ac1,OPER2)


      ASSIGN     2,PH3,EXP1,PF
                       specify machining time
      ADVANCE    PF2  advance machining time
*     bcall      &aadvanc(pf16,pf2,2)


      PRIORITY   1,BUFFER  set priority, get
*     bcall      &benter(pf16,ac1,OPER2)
      ENTER      OPER2     next free operator
*     bcall      &aenter(pf16,ac1,OPER2)

      ASSIGN     2,PH4,EXP1,PF
                             specify end time
      ADVANCE    PF2         advance end time
*     bcall      &aadvanc(pf16,pf2,3)

      LEAVE      OPER2       return operator
*     bcall      &aleave(pf16,ac1,OPER2)

      RELEASE    VTL1        release machine
*     bcall      &arelease(pf16,ac1,VTL1)
                  .
                  .
```

FIGURE 3

4. RESULTS OF SENSE AND GPSS/H MODEL OF FLMS

4.1 SENSE PREDICTIONS FOR FLMS MODEL

The above system was successfully coded and instrumented with SENSE. We now display some typical outputs of SENSE.

We show in Figure 4 the sensitivity of system FT due to changes in setup time, machining (process) time, and moving between resources (other) time for all resource groups (i.e. the VTLs, the polishers, etc.)

```
****  Current time   ******  1200000  ****

    SENSE FLOW LINE No. 4.05

***  Effect of MACHINE GROUP PERFORMANCE on
       --  TOTAL -- FLOW TIMES  ***

 % change in FLOW TIME due to 1% change in

OPN. at  *  setup  process  other * Total
MACH.GRP *                        *
*****************************************
    1    * .290    .070     .009  * .369
    2    * .225    .030     .016  * .271
    3    * .086    .043     .018  * .147
    4    * .108    .025     .011  * .144
    5    * .061    .017     .005  * .084
    6    * .165    .021     .009  * .195
*****************************************
```

Figure 4

This table can be used to discuss tradeoffs and the effects of a number of changes to the system. For example if we decreased all of the setups at machine group 2 (manual VLTs) by x% (e.g. 10%) we would change the average flow time and work-in-process by x*.225% (e.g. 10 * .225 = 2.25%). Likewise we could tradeoff the process and setup times at one machine group, increasing one and decreasing the other. Also, we could see the benefits of moving an operation from a machine group with large sensitivities to a machine group with low sensitivity. To produce this table by straight simulation we would have to run an additional 18 experiments (one for each entry). Thus the power of SENSE is very clear.

Figure 5 shows that percent change in system flow time (FT) due to unit change in lot size of each part. However in order to keep the annual output constant, we evaluate this sensitivity with a concurrent change in the arrival rate of the part type. This is an example of the ease of adapting or customizing the output of SENSE. Actually, two sensitivities (with respect to arrival rate and lot size) are evaluated for the system by SENSE and then combined to give the required sensitivity.

```
****  Current time   ******  1200000  ****

    SENSE FLOW LINE No. 4.05

***  Effect of PART INTRODUCTION RATE on
      --  TOTAL -- FLOW TIMES  ***

% change in FLOW TIME due to UNIT CHANGE in
  LOT SIZE, decreasing ARRIVAL RATE so
  ANNUAL PRODUCTION is CONSTANT


PART    * LOT *  EFFECT    * PART
NUMBER * SIZE* FLOW TIME * DESCRIPTION
*****************************************
   1    *  8  * -.500      * part number 1234
   2    *  7  * -.690      * part number abcd
   3    *  6  * -.423      * part number 5678
   4    *  7  * -.332      * part number wxyz
*****************************************
```

Figure 5.

Figure 6 illustrates the percent change in average flow time FT of all part types due to addition of one additional AGV. SENSE can also calculate the sensitivity with respect to the number of machines and operators.

```
****  Current time   ******  1200000  ****

    SENSE FLOW LINE No. 4.05

***  Effect of ADDITION of one RESOURCE on
      --  TOTAL -- FLOW TIMES

% change in FLOW TIME due to ADDITION of
  one UNIT to RESOURCE GROUP

Resource  *   EFFECT on     * GROUP
GROUP     *   FLOW TIME     * DESCRIPTION
*****************************************
    9     *    -0.877       * AGVs
*****************************************
```

Figure 6.

4.2 VALIDATION OF PREDICTIONS BY SUBSEQUENT EXPERIMENTS

As validation of the predictive power of SENSE, six more simulation runs of the GPSS/H code for the FLMS were performed. For each run one system parameter, such as lot size, number of AGVs, and setup time, etc., was changed, and the resultant system FT was observed and compared with the prediction by SENSE. Figure 7 display the results which are typical of the many experiments the authors have performed on all types of discrete event networks.

In the first experiment we increased the set-up time at machine group 1 (N/C VTL). SENSE had reported a sensitivity of .290 so that the average flow time after the +10% change should increase by 2.90%. We then changed the setup time in the model and reran the simulation. The predicted and actual average FT and the changes in the Flow Time are reported in the table above.

Entity changed	AVERAGE PART FLOW TIME SENSE pred.	Actual	SENSE change in Flow Time	Actual change in Flow Time
Original	-	8444	-	-
+10% setup MACH.GR. 1	8624	8610	180	166
+10% proc. MACH.GR. 1	8498	8492	54	48
+5% Arr. all Parts	8620	8637	176	193
+1 in Lot for Part 1	8477	8490	33	46
+1 in Lot -14% Arr. for Part 2	8385	8398	-59	-42
1 AGV	8370	8357	-74	-87

Figure 7.

In the second experiment we changed the service time for every cutting operation at the first machine group (N/C VTLS).

In the third experiment we changed the entire schedule so that 5% more of all part types were produced in the year. It is possible to combine a number of the sensitivities reported into one prediction. The answer is less accurate and we are reporting derivatives which by definition add linearly. In reality the system is non-linear so that large predictions or combining many sensitivities into one prediction are less accurate.

In the fourth experiment we increase the lot size for the second part while leaving the interarrival time of the lot unchanged. This is a single change to the system. In the fifth experiment the lot size and the interarrival time are changed so that total annual production is constant. The prediction is calculated from two sensitivities: the arrival rate (e.g. expt 3) and the lot size (e.g. expt 4).

In the sixth experiment we see the effect of adding and AGV to the system and note just a small decrease to the overall flow time. This suggests that no AGV should be added and that the analyst should consider removing an AGV and rerunning the simulation.

In general those parameters with small sensitivities are areas where resources may be removed. Note that the sensitivities for the AGV move times (other column in figure 4) are very small. This would suggest that one could substitute a slow less costly AGV system with little effects on the Flow Time.

SENSE identifies those parameters are critical and thus informs the analyst that these must be the more accurately estimated values.

5. CONCLUSION

Simulation is a powerful tool which is used in the analysis of systems. However it is expensive in terms of designer, programmer and computer time. Perturbation Analysis is method of getting more information out of each simulation run. SENSE is a tool which implements perturbation analysis. In the GPSS/H model of a flow line manufacturing system we showed how SENSE is used and the information it provides to an analyst and designer. SENSE not only identifies those areas where sensitivities are large (bottlenecks) but also identifies those areas where resources are less critical. The tool is an accurate and cost effective method of generating sensitivities of a model with respect to a wide variety of parameters.

REFERENCES

[1] See several articles in SME Journal of Manufacturing Systems, Vol.1, Nos. 1-2, 1983.

[2] Suri, R. and C.K. Whitney,"Decision Support Requirements in Flexible Manufacturing", SME Journal of Manufacturing Systems, Vol.3, No.1, 61-69, 1984.

[3] Suri, R. "Quantitative Techniques for Robotic Systems Analysis", chapter in Handbook of Industrial Robotics, S.Y. Nof (ed.), John Wiley 1985.

[4] Suri, R., and Diehl, G.W. "MANUPLAN: A Precursor to Simulation for Complex Manufacturing Systems", Proceedings of the Winter Simulation Conference, December 1985 (this volume).

[5] Suri, R. and J.W. Dille, "A technique of On-Line Sensitivity Analysis of Flexible Manufacturing Systems", to appear in Annals of Operations Research.

[6] Ho, Y.C., M.A. Eyler, and T.T. Chien, "A Gradient Technique for General Buffer Storage Design in a Serial Production Line", Int. J. Prod. Res. 17, 6, 1979, 557-580.

[7] Ho, Y.C., R. Suri, X.R. Cao, G.W. Diehl, J.W. Dille and M.A. Zazanis, "Optimization of Large Multiclass (Non-Product-Form) Queueing Networks Using Perturbation Analysis", Large Scale Systems 7, 165-180, 1984.

[8] Ho, Y.C., and X. Cao,"Perturbation Analysis and Optimization of Queueing Networks", J. Optim. Theory Applic. 40, 4, 559-581, August 1983.

[9] Suri, R., "Implementation of Sensitivity Calculations on a Monte Carlo Experiment", J. Optim. Theory Applic. 40, 4, 625-630, August 1983.

[10] Suri, R. and Cao, X., "The Phantom Customer and Marked Customer: Methods for Optimization of Closed Queueing Networks with Blocking and General Service Times," ACM Performance Evaluation Review, August 1983, 243-256.

[11] Suri, R., "Infinitesimal Perturbation Analysis of Discrete Event Dynamic Systems: A General Theory", Proc. 22nd IEEE Conf. Dec. and Control, Dec. 1983.

[12] Ho, Y.C., Eyler, M.A., and Chien, T.T., "A New Approach to Determine Parameter Sensitivities of Transfer Lines," Management Science 29, 6, June 1983, 700-714.

[13] Ho, Y.C. and X. Cao, "Perturbation Analysis of Sojourn Time in Queueing Networks", to appear in Operations Research.

[14] Zazanis, M.A. and R. Suri, "Comparison of Perturbation Analysis with Conventional Sensitivity Estimates for Regenerative Stochastic Systems", Operations Research, submitted January 1985.

[15] Woodside, C.M., "Response Time Sensitivity Measurement for Computer Systems and General Closed Queueing Networks", Performance Evaluation, Vol.4, 199-210, 1984.

[16] Suri, R. and Zazanis, M., "Perturbation Analysis Gives Strongly Consistent Estimates for the M/G/1 Queue", Management Science, submitted Sept. 1984.

Rajan Suri is currently Associate Professor of Industrial Engineering at the University of Wisconsin-Madison. He received his Bachelors degree in 1974 from Cambridge University (England) and his M.S. (1975) and Ph.d. (1978) from Harvard University. In 1981 he received the IEEE Donald P. Eckman award for outstanding contributions in his field. His current interests are in modelling and decision-support for manufacturing systems, specializing in Flexible Manufacturing Systems. He is the author of many journal publications and several books, and has chaired international conferences on this subject. He is a consultant to major industrial corporations, and also a principal of Network Dynamics, Inc., a Cambridge (MA) firm specializing in software for modelling manufacturing systems. His address is

Rajan Suri
University of Wisconsin-Madison
Dept. of Industrial Engineering
Madison, WI 53706
(608) 262-2686

Gregory W. Diehl is currently Technical Manager for Network Dynamics, Inc. He has worked for Bell Labs, Digital and BGS systems in performance analysis of Computer-Communication and Manufacturing Systems. He completed his A.B. (1979) at Oberlin College in Mathematics and Chemistry and S.M. (1980) and Ph.d. (1984) at Harvard in Engineering Sciences. He is the author of several publications in the area of discrete event systems modelling. His address is

Gregory Diehl
Network Dynamics, Inc.
1218 Massachusetts Ave.
Cambridge Mass. 02138
(617) 547-2036

Yu-Chi Ho is Gordon McKay Professor of Engineering and Applied Mathematics at Harvard University. He received his S.B. and S.M. degree in electrical engineering from M.I.T. and the Ph.D. degree in applied mathematics from Harvard. For the past twenty five years he has been working in the area of dynamic systems and control. He is on the editorial board of seven international scholastic journals and serves in various professional, administrative, and consulting capacities with government panels and professional societies. He holds four patents in numerically controlled machine tools, and has published 80 papers and two books. He first developed the idea of perturbation analysis of discrete event systems in 1978 and along with his colleagues Suri and Cao is a principal architect of the new technique. He is also a principal of Network Dynamics, Inc. His address is:

Pierce Hall
Harvard University
Cambridge, MA 02138
(617) 495-3992