

A CHTE PROJECT SUMMARY

Richard D. Sisson, Jr. and Aparna Varde

CHTE enhances the position of the heat treating industry by applying fundamental research to solve problems and to advance heat treatment technology. The research portfolio is member-driven, and the work enhances the industry's technology base, profitability, public image, and education of its members.

Data Mining in Heat Treatment

To model the thermal, mechanical, and metallurgical response of a part during quenching, it is necessary to know the heat transfer coefficient between the quenching fluid and the part as a function of temperature and position on the part. Currently, limited databases exist in several companies, laboratories, and CHTE. The determination of this heat transfer coefficient requires careful experiments designed to accurately measure temperature in a probe or part as a function of time during quenching. The data are then used to calculate the heat transfer coefficient as a function of temperature using an inverse mathematical method. These experiments are time consuming and costly.

Researchers at CHTE have developed data-mining techniques that can be used to extend this limited data. The goal of the project is to provide the user with the capability to input quenching process parameters to determine heat transfer coefficients.

From QuenchMiner to AutoDomainMine

Data obtained from experiments in heat treating offer a potential knowledge source that could be useful in making decisions. Knowledge must be identified from raw data and applied in such a way to assist in the decision making of a heat treater. This is the goal of the QuenchMiner™ system; to provide decision support in heat treating aimed at optimizing the processes involved. Users submit cases, and the system analyzes them using domain knowledge to output suggested decisions. (Note that this is only to support or assist users.)

The QuenchMiner logo shown in Fig. 1 summarizes the system's functions: the heat treater poses queries to the system; the system performs data mining and unlocks or discovers knowledge from raw data; the discovered knowledge is used to perform analysis similar to a domain expert; and the system conveys answers or suggestions to the user through a web interface.

The following example illustrates the decision support functionality of QuenchMiner. Figure 2 shows the user input for a particular situation. The user submits information on quenching conditions for the system to estimate the average heat transfer coefficients that would be obtained in the process. QuenchMiner estimates ranges of heat transfer coefficients in terms of categories such as high, medium and low, as shown in Fig. 3.

QuenchMiner only estimates ranges of parameters, such as cooling rates and heat transfer coefficients. In addition, it retrieves existing experimental data from heat transfer curves. However, users might be interested in estimating the actual heat transfer curve that would be obtained in an unperformed experiment given its input conditions. This motivated the development of a computational estimation technique that estimates the resulting graphs in heat treating experiments to save the time and resources that would be required to perform laboratory experiments. This esti-



Fig. 1—
QuenchMiner logo



Fig. 2 — Example of QuenchMiner estimation input



Fig. 3 — Example of QuenchMiner estimation output
 mation is based on analyzing the results of existing experimental data stored in a database.

The technique proposed to carry out the computational estimation at CHTE is called AutoDomainMine, which is based on data mining guided by the basic knowledge of the domain. The goals of AutoDomainMine with respect to quenching are:

- Estimate the resulting graph (that is; the heat transfer curve that would be obtained) given the input conditions of a quenching experiment.
- Estimate a set of input conditions that would produce the desired graph (that is, a heat transfer curve in a quenching experiment).

AutoDomainMine achieves these goals using a form of domain type-dependent data mining explained below. Data mining is the process of discovering interesting patterns and trends in large datasets to guide decisions about future activities. In AutoDomainMine, two data mining techniques of clustering and decision-tree classification are integrated into a learning strategy for estimation. AutoDomainMine first discovers knowledge from experimental results by integrating clustering and classification. It then uses the discovered knowledge to estimate curves resulting from new experiments given their input conditions.

Clustering is the process of placing a set of physical or abstract objects into groups of similar objects. Classification is a form of data analysis that can be used to extract models to predict categorical labels. These two

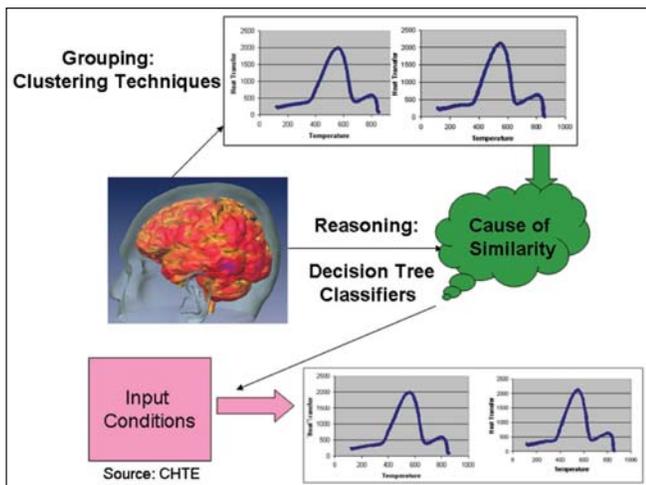


Fig. 4 — Automating typical learning strategies of scientists

High Temperature Thermocouples

- Accurate measurements to 4200°F (2200°C)
- Experts in Types R, S, B and W5
- Standard or built to your specs
- Delivered on time at competitive prices

For your free catalog, e-mail cwatkins@avureae.com or call 800-788-2732.

AVURE

Avure Autoclave Systems, Inc.
 (Formerly Flow Autoclave Systems, Inc.)
 3721 Corporate Drive
 Columbus, OH 43231
 614-891-2732
 Fax: 614-891-4568
 www.avureae.com

ISO-9001 Certified Quality System

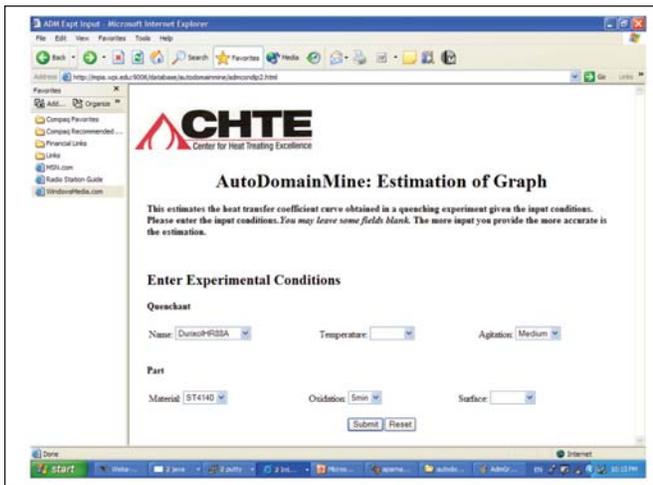


Fig. 5 — Input conditions to estimate graph (heat transfer curve)

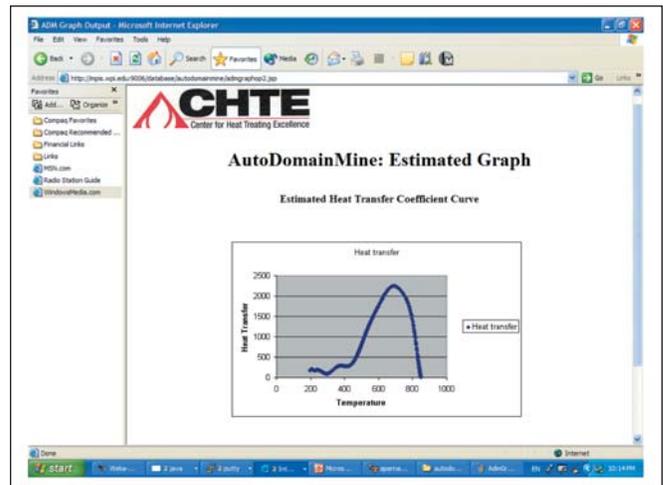


Fig. 6 — Estimated graph (heat transfer curve) for given input conditions

AutoDomainMine enhances decision support in heat treating.

data-mining techniques are integrated as illustrated in Fig. 4. Integration automates typical learning strategies of scientists in the targeted domains, and, therefore, is responsible for the domain type-dependent nature of AutoDomainMine. Scientists in domains such as heat treating often group experiments based on their results and then reason the causes of similarities between the groups. The grouping and reasoning processes are analogous to clustering and decision-tree classification, respectively.

The knowledge discovered from clustering and decision tree classifiers is then used to build a representative pair of input conditions and graph (heat transfer curve) per cluster. The combinations of input conditions leading to a cluster are identified through the decision tree paths. Therefore, when a new set of input conditions is submitted, the paths of the decision trees are traced to find the closest matching cluster. The representative graph of that cluster is the estimated graph in the experiment. Similarly, if a new heat transfer curve is submitted to estimate a set of conditions to obtain it, the given graph is compared with the representative graph to find the closest match. The corresponding set of input conditions is the set of estimated conditions to obtain the given heat transfer curve. Because the relative importance of the conditions has already been learned in the knowledge-discovery step, the estimation takes into account domain semantics.

Note that the knowledge-discovery process of AutoDomainMine is only a one-time step, which must be performed using the existing data in the database. The estimation process is a recurrent step, which must be performed each time the user submits a new experiment, so the time required for estimation is important. Therefore, in evaluating AutoDomainMine, two significant criteria are accuracy and efficiency. Accuracy indicates how close the estimation is to the real experiment, while efficiency refers to the time required to perform the estimation.

The AutoDomainMine system is evaluated using a test set of existing experiments distinct from the training set used for learning.

Several such tests are conducted and an example from the evaluation is presented below.

The user submits input conditions of a quenching experiment as shown in Fig. 5 to estimate the heat transfer curve that would be obtained. Figure 6 shows the estimated heat transfer curve. On comparing the estimated graph with the actual graph obtained in the laboratory experiment performed with the same input conditions (as stored in the test set), the domain experts conclude that the estimated result is similar to the real result. The response time of the system in performing the estimation is less than one minute, which is satisfactory for displaying at-a-glance information as confirmed by experts. Similar evaluations are also conducted to estimate conditions given the graph, and it is observed that the resulting estimation is accurate and efficient.

Likewise, on evaluating with several examples from the real data sets, it is observed that AutoDomainMine depicts good performance in terms of both accuracy and efficiency. Thus, it is a useful technique for computational estimation in the heat treating domain.

AutoDomainMine has several applications in heat treating. The main application is enhancing the decision support functionality of the earlier QuenchMiner system. While QuenchMiner estimates ranges of parameters for unperformed experiments and retrieves the heat transfer curves obtained from performed experiments, AutoDomainMine estimates the heat transfer curves that would be obtained in unperformed experiments, which goes a step beyond QuenchMiner to assist users by estimating more precise information helpful in making decisions about the corresponding real processes. The AutoDomainMine estimation can be used to select process parameters for quenching in the industry. Other applications include serving as the input to software tools, such as DANTE, DEFORM, SYSWELD, and CHT-bf, that perform simulations. In addition, Auto-DomainMine could be useful for intelligent tutoring systems in heat treating.

