ABSTRACTS-YEAR 2000

DISSERTATION

COMPUTER SCIENCE

Diss CS-00-1
TRAFFIC MODELING FOR WWW
Panjai Tantatsanawong
Dr. Kanchana Kanchanasut
An analytical modeling approach is introduced in this research to capture actual characteristics of the WWW traffic. The model consists of two submodels, one for the client side and the other for server side (WWW server or proxy cache server). The workload on the client side can be characterized by type of the requested pages (HTML, inlined image and CGI) and on the server side by the requested file types. Zipf's law is applied to investigate the popularity of the workload classes. From workload characteristics of the WWW sites considered, the WWW server can be classified into two types: a static WWW server, which provides static information WWW pages and an interactive WWW server, which provides real time processing such as multi-tier or E-commerce applications. The client and the server models are synchronized with hourly time intervals. On the client side, several TCP performance effects are included in the model such as TCP connection setup (three-way handshake), persistent connection and slow-start algorithms.

On the server side, modeling and forecasting of hourly requested transactions are made using a combination of Box-Jenkins models with seasonality removal. Fourier analysis and standardization are used to render seasonal data into non-seasonal data. However, seasonality can not be completely removed by this approach, so both seasonal and non-seasonal models are employed and the best models are selected using posterior possibility criterion (PPC). In order to improve the forecasting accuracy, forecast values obtained by the developed models are adjusted by the difference percentage of the day of the week and by using this percentage adjusted values in place of the actual data, which are not available at the time of forecasting. With this adjustment, the developed models can provide very accurate forecast values, except for the case where the higher level of hierarchical proxy is shared by child proxies located in a different time zone.

Finally, multiple-class closed and opened queuing network models are developed to estimate response time at the client side and the server side respectively. The proposed model can satisfactory predict the impact on the client response time and the network utilization as a function of HTTP (version 1.0 and 1.1), implementing proxy cache, hierarchical proxy cache and improving the cache with a prefetch mechanism.

Diss CS-00-2
BACKPROPAGATION NETWORKS FOR FORECASTING
Songyot Sureerattanan
Prof. Huynh Ngoc Phien
In recent years, artificial neural networks (ANNs) have been extensively used in various fields. Among them, backpropagation (BP) networks appear to be most popular and have been widely used in many applications such as forecasting and classification. To predict the future outcome values with an acceptable level of accuracy, a BP network has to be trained with a large sample of historical data that have been collected over a given time period. The BP network will then learn to generalize and extrapolate from new data to predict the future outcomes. However, there have been several problems encountered. One is how to determine the appropriate structure of the network for a particular problem. Another problem is its slow convergence (and no convergence in some cases), so that many iterations are required to train even a simple network.

The structure of the network seriously affects the performance of the network model. As the network becomes more complex, the training time will increase. Therefore, the network should be kept as simple as possible. As the number of nodes in the input and output layers are application-dependent, the remaining problems are how to optimally choose the number of hidden layers and the number of hidden nodes. For many applications, they are determined by trial-and-error. Generally, when the number of parameters (the number of weights and biases) increases, the mean squared error (MSE) should be reduced. Therefore, it is difficult to determine the best network model by using only MSE. Instead, the Bayesian Information Criterion (BIC) was proposed in this study to select the best model from the candidate models having different numbers of parameters. It should be noted that the BIC penalizes the model for having more parameters and therefore tends to result in a smaller model. A new stopping rule was proposed to systematically determine the appropriate network structure using a procedure that gradually increases the network complexity until the current value of BIC is greater than the previous one or the decrease in the value of BIC becomes very small.

Two new algorithms were devised to speed up the convergence of BP networks:
· The first proposed algorithm was obtained by applying the adaptive neural model with the temperature momentum term to the Kalman filter (KF) with the momentum term.
· For advanced refinement, the nonlinear neural network problem can be partitioned into the nonlinear part in the weights of the hidden layers and the linear part in the weights of the output layer. By employing the conjugate gradient method for the nonlinear part and the KF algorithm for the linear part, we arrived at the second proposed algorithm. After the weights of the hidden layers are obtained by using the conjugate gradient method, the weights of the output layer (in a linear problem) are readily solved by KF. The partition allows the nonlinear and linear parts of the search to be conducted in a reduced dimensional space, resulting in acceleration of the training process. Consequently, the second proposed algorithm can greatly improve the

convergence speed. From simulation experiments with three data sets; namely, daily stream flow (rainfall-runoff) data, quarterly data on exports and gross domestic product (GDP) of Thailand, and daily data on stock prices in Thai market, it was found that the BIC and these algorithms could perform satisfactorily in all cases considered.

The BIC criterion and the two algorithms were introduced without any conditions. Consequently, they should be generally applicable to any type of data.

Keywords: backpropagation networks, forecasting, network structure, convergence rate, Bayesian Information Criterion, Kalman filter, adaptive neural model, conjugate gradient.

Diss CS-00-3

MODELING AND IMPLEMENTATION OF THAI HAND WRITING CHARACTERS RECOGNITION

Chomptip Pornpanomchai
Dr. Dencho N. Batanov

An automatic handwritten character recognition system is one of the most difficult tasks in image processing. A handwritten character recognition task has two main problems, an unlimited handwritten character style and an uncertain handwritten character geometric shape. This thesis uses a feature-based, fuzzy logic and object-oriented approach (FBFLOOA) to recognize Thai handwritten characters. A feature-based approach is used to extract Thai handwritten character features. Fuzzy logic technology is used to justify the uncertain geometric shape of a handwritten character. An object-oriented approach is used to analyzed and design a Thai handwritten character recognition model.
The study concludes that all Thai handwritten characters consist of three basic geometric shapes or basis sub form features, a line, an arc and a circle. These three basic sub form features are connected with an appropriate coordination or precision feature to construct a Thai character. A basic sub form feature is a terminal node of a form feature hierarchy that inherits some properties from their super-classes. In terms of form feature hierarchy and properties inheritance, it implies that an object-oriented approach is an effective approach that can analyze and design a Thai handwritten character recognition model. Moreover, the study uses the Unified Modeling Language (UML) standard diagrams, which are 1) a use case diagram, 2) class-relationship diagram, 3) dynamic diagram, 4) activity diagram, 5) sequence diagram and 6) collaborate diagram to represent a Thai handwritten character model.

The study tested five hundred Thai handwritten words, which consist of four thousand one hundred and fifty one characters. The experiment was divided into two phases. The first phase used only the FBFLOOA to recognize a handwritten character and the second phase used the FBFLOOA combined with a Thai dictionary file to search for a correct answer for a non-recognition character. The first phase experimental results demonstrated a recognition accuracy of 91.21%, 7.51 % non-recognition and 1.28 % mismatch. The second phase precision results were 98.21%, 0.51% non-recognition and 1.28% mismatch. On average, both phases had a recognition speed of 6.71 seconds per character. The FBFLOOA executed program size was 189 KB and the Thai dictionary file was 853 KB. The small size of Thai handwritten character recognition program is suitable for portable computers, non-keyboard calculators and handheld computers.

Based on the results, this study concludes that the FBFLOOA is an effective handwritten character recognition approach, which can be applied to identify not only Thai handwritten characters but also the handwritten characters of any language because all language characters consist of form and precision features.

Diss CS-00-4

RELATIONSHIPS BETWEEN BEZIER AND BALL BASES, AND INTERSECTION ALGORITHM FOR BEZIER PATCHES

Nattawit Dejdumrong
Prof. Huynh Ngoc Phien

The need for the determination of Surface Intersection exists in many real-world applications. Such determination is based on a recurrent operation so that its computation should be fast, reliable and suitable for the surfaces involved. Two methods are applied and two new methods are proposed in this work for determining the intersection of two Bezier surfaces: Subdivision, Marching Methods, Polyhedron Intersection, and Hybrid between subdivision and polyhedron intersection. Since Bezier and two generalized Ball surfaces can be treated as polynomials, all these methods can be properly applied to the calculation of the intersection of two generalized Ball surfaces by converting them to the Bezier surfaces. However, the relationships between Bezier and Said-Ball (or Wang-Ball) surfaces must be explicitly defined. Using homogeneous coordinates in the de Casteljau algorithm for the points on an polynomial Bezier surface, the de Casteljau algorithm to compute the points on a rational Bezier surface is readily obtained. Modifying some parts of four methods, algorithms for determining the intersection of two rational Bezier surfaces are eventually achieved. Similar to the case of rational surfaces, the relationships between Bezier and Said-Ball (or Wang-Ball) surfaces must be appropriately introduced in order to make use of the intersection methods for Bezier surfaces.

Subdivision methods reduce the problem of surface/surface intersection to the intersection of two planes by recursively subdividing two surfaces until they are relatively flat. Bounding boxes of the surfaces are used to coarsely locate the intersection parts. The irrelevant parts that do not partcipate in the intersection will be gradually eliminated. Only relevant pairs of subsurfaces will be subdivided until they meet a predetermined tolerance. Ultimately, the final result is a collection of line segments.

An alternative method is to roughly locate the intersection part by the intersection of two polyhedra. Instead of directly

calculating the intersection from the surfaces, the polyhedra with quadrilateral facets that are readily formed by the control nets are considered. By this approach, the problem of surface/surface intersection can be reduced to the problem of polyhedron/polyhedron intersection. Then the problem can be further simplified by triangle/triangle intersection. The result is one or more sets of consecutive line segments, which will be used as a starting set for further refinement to obtain the true intersection.
Combining the two previous methods (subdivision and polyhedron intersection), a hybrid method is obtained. It integrates the subdivision method with the intersection of two triangulated polyhedra. Thus, the final result is much more precise than that for plane/plane intersection. Consequently, it is closer to the exact result.

Unfortunately, none of the three methods can produce the true intersection curves. Further refinement of the intermediate result needs to be performed. Two marching methods, using Tangential or Circular steps, are used to calculate exact intersection points. A major difference between the two techniques is the step size. In the method using tangential steps, the step size has to be fixed and predefined while that of the circular step is dynamic and automatically changed. The selection of which techniques to be used depends on the problem itself: if the result is crooked or winding, the circular step is recommended, otherwise it is sufficient to use the tangential step.

In this study, all of the methods focus on the intersection of two Bézier surfaces. To make use of these methods for the Said-Ball and Wang-Ball surfaces, conversions from these Ball surfaces into Bézier surfaces are necessary. This dissertation presents relationships among Bézier, Said-Ball and Wang-Ball curves. These are used as the foundations for the relationships between Bézier and Ball surfaces. Using the relationships between Bézier and Ball surfaces, the intersection of two Ball surfaces is readily obtained by converting Ball surfaces into Bézier surfaces and determining the intersection of two Bézier surfaces. Thus, this is adequate for determining the intersection of two different kinds of surfaces.

Furthermore, all these methods can be applied to the intersection of two rational Bézier surfaces. Using the de Casteljau algorithm for a rational Bézier surface, the proposed methods based on subdivision for Bézier surfaces can readily be extended to rational Bézier surfaces. The other methods can be readily adapted to the rational Bézier surfaces. Similarly, using the relationships between rational Bézier and rational Ball surfaces, the intersection of two rational Ball surfaces is readily obtained by converting rational Ball surfaces into rational Bézier surfaces and determining the intersection of the two resulting rational Bézier surfaces.