Economides, A.A.: **Learning automata for multimedia traffic control.** *Proceedings International Conference on Circuits, Systems & Computers '96*, pp. 555-560, 1996.

LEARNING AUTOMATA FOR MULTIMEDIA TRAFFIC CONTROL

Anastasios A. Economides

University of Macedonia
Thessaloniki 54006, GREECE

Tel # +3031-891799
Fax # +3031-891750
economid@macedonia.uom.gr

Abstract: Learning automata are used at the source nodes of a multimedia network to dynamically control the externally arriving traffic to the network. A new fixed structure learning automaton algorithm is used to control the traffic.

At every network source, each different traffic type (data, voice, video, e.t.c.) is controlled by a different learning automaton. Every learning automaton makes its decisions based on information about the current performance of its associated traffic.

Key-words: multimedia networks, learning automata, traffic control, traffic management, adaptive admission control, adaptive routing.

1. INTRODUCTION

Advances in data transmission and switching over the last years promise deployment of communication systems with bandwidth and switching speeds that are orders of magnitude higher than the current systems. However, the commercial relevance of a networking technology depends not only on the functional aspects, but even more on the cost of introducing it into the existing infrastructures through end-system integration, cabling, interconnection and management [1].

Video on demand, videoconferencing, multimedia database retrieval, digital libraries, tele-education, tele-medicine, multiparty interactive video games and distributed multimedia cooperative work are just a few of the many distributed multimedia applications that have emerged in recent years. Widespread use of such applications will require networks that can offer transport services to any type of media. Communications services have traditionally been provided by different types of networks: telephone networks for voice, cable TV networks for video, data networks for computer communications. In contrast, multimedia networks must integrate communication services for all media types in a single network. It is widely accepted that multimedia networks should be based on a packet (or cell) switching paradigm [2].

Multimedia networks will offer efficient statistical multiplexing of multiple traffic types (data, voice, video etc.) on a single high speed integrated services network. Because of the real-time characteristics of audio/visual applications, multimedia networks must guarantee stringent Quality Of Service (QOS). Multimedia applications are not only bandwidth-hungry, bur require also to meet real-time delay constraints, delay "jitter" restrictions, packet loss guarantees etc.

When multimedia communications are handled over a network, a severe performance bottleneck and service quality degradation is possible at the switches. To solve this problem, dynamic traffic control to adjust to current network conditions is necessary. Efficient traffic control can make orders of magnitude difference in terms of cost-effectiveness.

When a multimedia application requests communication through the network, an admission control decision should be made. If there are not available network resources, the connection request is rejected; otherwise the connection is established and a path is selected for transfering the traffic.

In this paper, we present a dynamic traffic control scheme based on learning automata algorithms. These are adaptive control algorithms for highly uncertain systems [3]. They select an action and then update their action decisions according to the outcome of the selected action. The greatest potential of the learning automata methodology is that it permits the control of very complex dynamic systems. Even when little information is available, they act to minimize the effects of future system changes. Learning automata have been applied to routing problems like telephone routing [4-7], datagram routing [8,9], and virtual circuit routing [10,11]. British Telecom has already developed a learning automata routing system for use in their long distance network [12].

In section 2, we introduce a new fixed structure learning automaton algorithm, that has fast rate of convergence and is suitable for switching environments, eg. multimedia networks. In section 3, we propose that learning automata at the source nodes of multimedia networks control the externally arriving traffic. We suggest that at each source, there is a different learning automaton for each traffic type between each specific source-destination. In section 4, we discuss candidate performance metrics that may be used by the learning automata in controlling the traffic. Finally, in section 4, we conclude on the proposed approach.

2. A NEW LEARNING AUTOMATON ALGORITHM

Here, we present the *STack ARchitecture* (STAR) automaton, a fixed structure, multi-action, reward-penalty learning automaton. STAR is characterized by the star shaped structure of its state transition diagram. There are r branches, one for each action in the action set. There is an additional, neutral state, from which each action is selected equiprobably.

At each instant n, the automaton selects probabilistically (according to the action probability vector $p(n)$) an action $a(n) = i$ from the finite action set a. The probability that the automaton selects action i, at time n is the action probability $p_i(n) = Prob[a(n) = i]$; we have $\sum_{i=1}^{r} p_i(n) = 1 \; \forall n$. The environment responds with $\beta(n)$; when the response is favorable (reward) $\beta(n)$=0, when it is unfavorable (penalty) $\beta(n) = 1$.

The automaton can be in any of $r+1$ states, $\mathbf{\Phi} = \{0, 1, ..., r\}$ (Figure). When the automaton is in state i ($i = 1, 2, ..., r$), it performs action i with probability close to 1 and any other action with very small probability (close to 0).

On the other hand, the state 0 is a special, so-called "neutral" state: when in that state, the automaton choooses any of the r actions equiprobably.

The transitions from state to state are described by the following rules, where $i, j = 1, ..., r$ and $j \neq i$.

1. When in state 0 and chosen action is i, if rewarded go to state i w.p. $1 - \hat{\delta}_1$ or stay in state 0 w.p. $\hat{\delta}_1$:

   $F_{0i0,i} = 1 - \hat{\delta}_1, \;\; F_{0i0,0} = \hat{\delta}_1, \;\; F_{0i0,j} = 0$

   but if punished, go to state i w.p. $\hat{\delta}_2$, or stay in state 0 w.p. $1 - \hat{\delta}_2$:

   $F_{0i1,i} = \hat{\delta}_2, \;\; F_{0i1,0} = 1 - \hat{\delta}_2, \;\; F_{0i1,j} = 0$

2. When in state i, $i \neq 0$ and chosen action is i, if rewarded stay in state i w.p. $1 - \delta_1$ or go to state 0 w.p. δ_1:

   $F_{ii0,i} = 1 - \delta_1, \;\; F_{ii0,0} = \delta_1, \;\; F_{ii0,j} = 0$

   if punished, go to state 0 w.p. $1 - \delta_2$, or stay in state i w.p. δ_2:

   $F_{ii1,i} = \delta_2, \;\; F_{ii1,0} = 1 - \delta_2, \;\; F_{ii1,j} = 0$

3. When in state i, $i \neq 0$ and chosen action is j, if rewarded stay in state i w.p. δ_1 or go to state 0 w.p. $1 - \delta_1$:

   $F_{ii0,i} = \delta_1, \;\; F_{ii0,0} = 1 - \delta_1, \;\; F_{ii0,j} = 0$

   if punished, go to state 0 w.p. δ_2, or stay in state i w.p. $1 - \delta_2$:

   $F_{ii1,i} = 1 - \delta_2, \;\; F_{ii1,0} = \delta_2, \;\; F_{ii1,j} = 0$

In the next section, we use the STAR for controlling the various traffic streams in multimedia networks.

3. LEARNING TRAFFIC CONTROL

At every source node $[s.]$, there are many learning automata, each one for a particular traffic type (data, voice, video etc.) and destination node $[.d]$.

Each learning automaton rejects or routes newly arriving calls of its associated traffic type. Let there exist $r - 1$ paths between source-destination $[sd]$. The decisions of the automaton depend on the performance requirements and constraints of its associated traffic type and are based on the achieved performance.

The actions of each automaton are to route the call to its destination through a particular path i, $(i = 1, ..., r - 1)$, or to reject the call by sending it through a fictitious path r.

Next, we describe the operation of an automaton for the traffic type c between source-destination $[sd]$.

When the automaton is in state 0, it selects any path with equal probability $1/r$. Let the automaton select path $i, (i = 1, ..., r)$: if the performance of the traffic type c on this path is "good", the automaton goes to state i with large probability $1 - \hat{\delta}_1$ or stays in state 0 with small probability $\hat{\delta}_1$ $(0 \leq \hat{\delta}_1 \ll 1)$; otherwise, the automaton goes to state i with small probability $\hat{\delta}_2$ or stays in state 0 with large probability $1 - \hat{\delta}_2$ $(0 \leq \hat{\delta}_2 \ll 1)$.

When the automaton is in state i, $(i = 1, ..., r)$, it selects path i with probability $1 - \theta$, $(0 \leq \theta \ll 1)$, and every other path j, $(j = 1, ..., r, \;\; j \neq i)$ with probability $\theta/(r-1)$.

Let the automaton select path i: if the performance of the traffic type c on this path is "good", the automaton stays in state i with large probability $1 - \delta_1$ or goes to state 0 with small probability δ_1 $(0 \leq \delta_1 \ll 1)$; otherwise, the automaton stays in state i with small probability δ_2 or goes to state 0 with large probability $1 - \delta_2$ $(0 \leq \delta_2 \ll 1)$.

Let the automaton select path $j (j \neq i)$: if the performance of the traffic type c on this path is "good", the automaton stays in state i with small probability δ_1 or goes to state 0 with large probability $1 - \delta_1$; otherwise, the automaton stays in state i with large probability $1 - \delta_2$ or goes to state 0 with small probability δ_2.

Thus, summing up the main points of the routing and admission control, let a call of type c arrive at source $[s.]$ requiring transfer to destination $[.d]$. Based on its current action probabilities, assume that the corresponding automaton select path i for routing the call. Then, we measure the performance of traffic type c over the selected path i, as well as the performance of traffic type c over the other paths j. If the performance of the call meets its QOS requirements and is better than the performance of similar calls over the other paths, then we say that it was a "good" decision, otherwise it was a "bad" decision. For the special case of action r, we say that the decision to reject a call is "good", if the call can not meet its QOS requirements on any path. Subsequently, the automaton moves to the next state as described above. When a new call of type c arrives at source $[s.]$ destined to $[.d]$, the automaton selects a path according to the probabilities of its current state.

In the next section, we describe QOS requirements for multimedia traffic.

4. PERFORMANCE METRICS

Multimedia networks will carry many different multimedia applications and each one of them will communicate many different traffic types (data, voice, sound, video, image etc.). Each traffic type of a particular application has special QOS requirements that must be met. For example, HDTV requires cell loss probability $< 10^{-12}$ and cell delay $< 40 msec$. The tolerance regions of various data, voice and video traffic calls are described below:

voice: cell loss $< 10^{-4}$ to 10^{-6}
cell delay $< 10 msec$ to $150 msec$

data: cell loss $< 10^{-9}$ to 10^{-12}
cell delay $< 10 msec$ to $500 sec$

video: cell loss $< 10^{-9}$ to 10^{-12}
cell delay $< 10 msec$ to $10 sec$

Depending on the specific traffic type, the performance metric used in the learning automaton could be the virtual path length [11], the unfinished work [10], the cell delay [11], the cell delay jitter, the ratio of the lost (blocked) cells over the cells that are sent etc.

Let $X_i^c(k)$ be the measurements of traffic type c at path i of the performance metric used in the learning automaton. Define

$$\hat{\mu}_i^c(k+1) = (1-\epsilon)\hat{\mu}_i^c(k) + \epsilon X_i^c(k) \qquad \epsilon \in [0,1]$$

$$(\hat{X}_i^c(k+1))^2 = (1-\epsilon)(\hat{X}_i^c(k))^2 + \epsilon(X_i^c(k))^2$$

$$(\hat{\sigma_i^c}(k+1))^2 = (\hat{X}_i^c(k+1))^2 + (\hat{\mu}_i^c(k))^2$$

Summarizing, the automaton selects a path i for sending a call of type c. This action will be a "good" decision, if the performance metric of traffic type c at path i meets its QOS requirements and is better than the performance metric of traffic type c over the other paths:

$$\begin{array}{l} \hat{\mu}_i^c(k+1) < QOS_\mu^c \text{ and} \\ \hat{\mu}_i^c(k+1) = \min\limits_j \{\hat{\mu}_j^c(k+1)\} \\ \text{or} \\ (\hat{\sigma_i^c}(k+1))^2 < QOS_{\sigma^2}^c \text{ and} \\ (\hat{\sigma}_i^c(k+1))^2 = \min\limits_j \{(\hat{\sigma}_j^c(k+1))^2\} \end{array}$$

Then the automaton updates its state and select the path for the next call of type c.

5. CONCLUSIONS

In this paper, we use learning automata at the source nodes of a multimedia network to dynamically control the externally arriving multimedia traffic. Each learning automaton is associated with a particular traffic type and source-destination pair. Each automaton can be in any of $r+1$ states, where $r-1$ is the number of paths between the source-destination. It makes transitions from state to state depending on the achieved performance of its actions, which are to reject a new call or route it through a particular path. The automaton takes its actions according to fixed probabilities associated with its current state. If the selection of a path results in good performance for the particular traffic type, the automaton tends to stay in the state that patronage this path.

REFERENCES

1. H.J.Stuttgen, *Network evolution and multimedia communication*, IEEE Multimedia, pp. 42-59, Fall 1995.

2. J. Liebeherr, *Multimedia networks: issues and challenges*, IEEE Computer, pp. 68-69, April 1995.

3. K. Narendra and M.A.L. Thathacher, *Learning Automata: An Introduction*, Prentice Hall, 1989.

4. K.S. Narendra and E.A. Wright and L.G. Mason, *Application of learning automata to telephone traffic routing and control*, IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-7, No.11, pp. 785-792, Nov. 1977.

5. K.S. Narendra and M.A.L. Thathachar, *On the behavior of a learning automaton in a changing environment with application to telephone traffic routing*, IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-10, No. 5, May 1980.

6. P.R. Srikantakumar and K.S. Narendra, *A learning model for routing in telephone networks*, SIAM J. Control and Optimization, vol-20, no 1, Jan. 1982.

7. K.S. Narendra and P. Mars, *The use of learning algorithms in telephone traffic routing - a methodology*, Automatica, 19, (5), pp. 495-502, 1983.

8. M.S. Chrystall and P. Mars, *Adaptive routing in computer communication networks using learning automata*, Proc. of IEEE Nat. Telecomm. Conf., pp. A3.2.1-7, 1981.

9. A.A. Economides and J.A. Silvester, *Optimal routing in a network with unreliable links*, Proc. of IEEE Computer Networking Symposium, pp. 288-297, IEEE 1988.

10. A.A. Economides, P.A. Ioannou and J.A. Silvester *Decentralized adaptive routing for virtual circuit networks using stochastic learning automata*, Proc. of IEEE Infocom 88 Conference, pp. 613-622, IEEE 1988.

11. A. A. Economides, *Learning automata routeing in connection-oriented networks*, International Journal of Communication Systems, Vol. 8, pp. 225-237, 1995.

12. P. Chemouil, M. Lebourges and P. Gauthier, *Performance evaluation of adaptive traffic routing in a metropolitan network: a case study*, Proc. IEEE Globecom '89, pp. 314-318.

Author's biography: Anastasios A. Economides was born and grew up in Thessaloniki, Greece. He received the Diploma degree in Electrical Engineering from Aristotle University of Thessaloniki, in 1984. After receiving a Fulbright and a Greek State Fellowship (1985), he continued for graduate studies at the United States. He received a M.Sc. and a Ph.D. degree in Computer Engineering from the University of Southern California, Los Angeles, in 1987 and 1990, respectively. During his graduate studies, he was a research assistant performing research on routing and congestion control. He is currently an Assistant Professor of Computer Networks at the University of Macedonia, Thessaloniki. His research interests are in the area of Performance Evaluation, Design, Optimization and Control of High-Speed Networks, Multimedia Networks, B-ISDN, ATM. He is a member of IEEE, ACM, INFORMS, ΕΠΥ, TEE.