

## Chapter XIII

# Information–Theoretic Methods for Prediction in the Wireless and Wired Web

**Dimitrios Katsaros**

*Aristotle University of Thessaloniki, Greece*

### ABSTRACT

Discrete sequence modeling and prediction is an important goal and challenge for Web environments, both wired and wireless. Web clients' data-request forecasting and mobile location tracking in wireless cellular networks are characteristic application areas of sequence prediction in such environments. Accurate data-request prediction results in effective data prefetching, which combined with a caching mechanism can reduce user-perceived latencies as well as server and network loads. Also, effective solutions to the mobility tracking and prediction problem can reduce the update and paging costs, freeing the network from excessive signaling traffic. Therefore, sequence prediction comprises a very important study and development area. This chapter presents information-theoretic techniques for discrete sequence prediction. It surveys, classifies, and compares the state-of-the-art solutions, suggesting routes for further research by discussing the critical

issues and challenges of prediction in wired and wireless networks.

### INTRODUCTION

The proliferation of wireless cellular networks and the penetration of Internet services are changing many aspects of Web computing. Constantly increasing client populations utilize diverse devices to access the wired and wireless medium, and various heterogeneous applications (e.g., traffic- and weather-condition broadcasting, streaming video) are being developed to satisfy the eager requirements of the clients. In this environment, seamless and ubiquitous connectivity as well as low client-perceived latencies are two fundamental goals. The first goal calls for smart techniques for determining the current and future location of a mobile node, and the second goal calls for efficient and effective techniques for deducing future client requests for information pieces.

Both of the aforementioned problems are related to the ability of the underlying network to record, learn, and subsequently predict the mobile user's behavior, that is, its movements or its information needs. The success of the prediction is presupposed and is boosted by the fact that mobile users exhibit some degree of regularity in their movement and/or in their access patterns (Bhattacharya & Das, 2002; Nanopoulos, Katsaros, & Manolopoulos, 2003). This regularity may be apparent in the behavior of each individual client or in client groups. The detection of regularity patterns can lead to drastic improvements on the underlying wireless network's performance. Accurate data-request prediction results in effective data prefetching (Nanopoulos et al., 2003) combined with a caching mechanism (Katsaros & Manolopoulos, 2004; Vakali, 2001) can reduce user-perceived latencies as well as server and network loads. Also, effective solutions to the mobility tracking and prediction problem can reduce the update and paging costs, freeing the network from excessive signaling traffic (Bhattacharya & Das, 2002).

These issues had been treated in isolation, but pioneering works (Bhattacharya & Das, 2002; Vitter & Krishnan, 1996) are paving the way for treating both problems in a homogeneous fashion. They exhibited the possibility of using methods that have traditionally been used for data compression (thus characterized as information-theoretic) in carrying out prediction. The unifying principle is that they model the respective state space as finite alphabets comprised of discrete symbols. In the mobility tracking scenario, the alphabet consists of all possible sites (cells) where the client has ever visited or might visit (assuming that the number of cells in the coverage area is finite). In the request-prediction scenario, the alphabet consists of all the data objects requested by the client plus the objects that might be requested in the future (assuming that the objects come from a database and thus their number is finite).

A smart network can record the movement (request) history and then construct a mobility (data-access) model for its clients. The history refers to the past, but the model is probabilistic and extends to the future. As uncertainty is inherent in mobile movements or requests, we can consider the requests to be the outcome of an underlying stochastic process, which can be modeled using established information-theoretic concepts and tools (Misra, Roy, & Das, 2004; Vitter & Krishnan, 1996).

In our earlier work, reported in Nanopoulos et al. (2003), we described a framework that was able to embrace some algorithms that had been presented in the context of Web prefetching. That framework was able to present those algorithms as variations of the standard PPM (prediction by partial match) technique. The emphasis of the framework was on differentiating between the techniques based on whether they record contiguous subsequences or noncontiguous subsequences. From that work, it was clear that the discovery of noncontiguous subsequences required considerable computational effort and could be realized only through off-line algorithms.

Extending this work, this chapter provides a unifying framework for all the methods, which deal with the issues of location tracking and prediction and request forecasting using known information-theoretic structures, not only the PPM structures; the framework treats them as (variable or fixed-length) Markov chains and presents the different families of methods, categorizing the state-of-the-art algorithms into their respective families. It mainly deals with the discovery of contiguous subsequences, although it can relatively easily be extended to include noncontiguous subsequences. An important objective of the chapter is to include in the presentation not only the algorithms that are familiar in the wireless-communications community, but also techniques that have been developed in other disciplines, like computational biology, machine learning, and the

14 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: [www.igi-global.com/chapter/information-theoretic-methods-prediction-wireless/28721](http://www.igi-global.com/chapter/information-theoretic-methods-prediction-wireless/28721)

## Related Content

---

### Configuring Methods for Deadlock-Free Routing

Valentin Olenev and Alexandr Karandashev (2022). *International Journal of Embedded and Real-Time Communication Systems* (pp. 1-20).

[www.irma-international.org/article/configuring-methods-for-deadlock-free-routing/289199](http://www.irma-international.org/article/configuring-methods-for-deadlock-free-routing/289199)

### On Strengthening of Weak Algebraic Manipulation Detection Codes

Maksim Alekseev (2015). *International Journal of Embedded and Real-Time Communication Systems* (pp. 1-26).

[www.irma-international.org/article/on-strengthening-of-weak-algebraic-manipulation-detection-codes/168514](http://www.irma-international.org/article/on-strengthening-of-weak-algebraic-manipulation-detection-codes/168514)

### Requirements Traceability within Model-Based Testing: Applying Path Fragments and Temporal Logic

Vanessa Grosch (2013). *Adoption and Optimization of Embedded and Real-Time Communication Systems* (pp. 88-107).

[www.irma-international.org/chapter/requirements-traceability-within-model-based/74243](http://www.irma-international.org/chapter/requirements-traceability-within-model-based/74243)

### Incompleteness and Unpredictability of Networked Communications in Use

Dixi Louise Strand (2007). *Designing for Networked Communications: Strategies and Development* (pp. 26-51).

[www.irma-international.org/chapter/incompleteness-unpredictability-networked-communications-use/8191](http://www.irma-international.org/chapter/incompleteness-unpredictability-networked-communications-use/8191)

### Privacy-Aware Contact Sharing for Groups in P2P Networks

Tommi Meskanen, Jarkko Kuusijärvi and Valtteri Niemi (2021). *International Journal of Embedded and Real-Time Communication Systems* (pp. 36-54).

[www.irma-international.org/article/privacy-aware-contact-sharing-for-groups-in-p2p-networks/291966](http://www.irma-international.org/article/privacy-aware-contact-sharing-for-groups-in-p2p-networks/291966)