Information Sciences & Technology

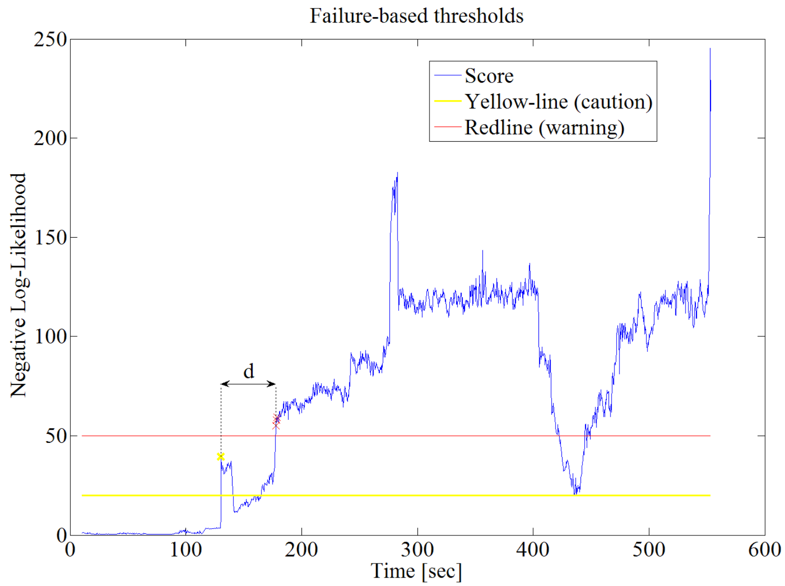
Ames Research Center

Approximations to Optimal Alarm Systems for Anomaly Detection

Rodney A. Martin, Ph.D.
NASA Ames Research Center
October 30, 2008

Background

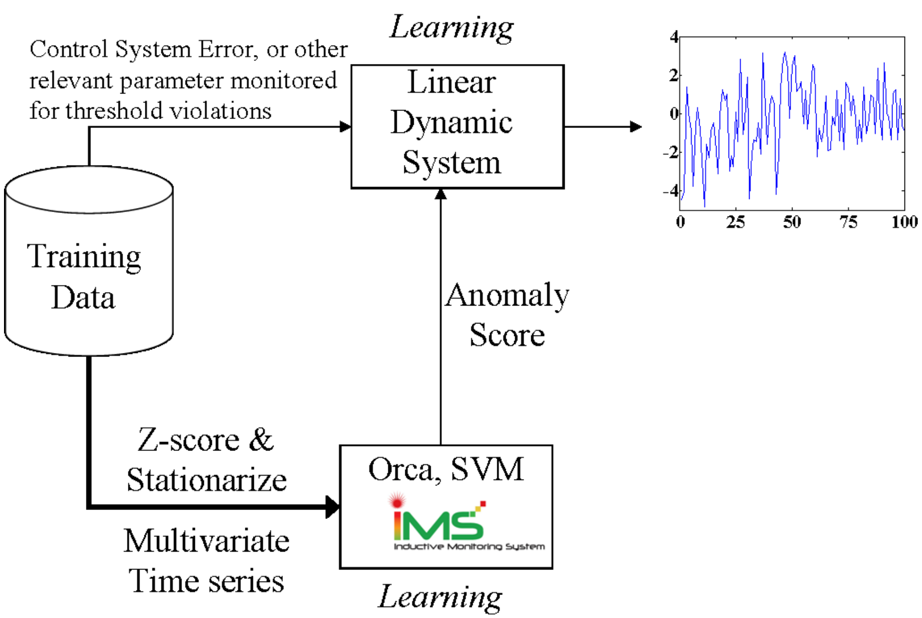
- State of the art involves setting a threshold ("redline") based upon physics of failure
  - Butas et al. (J-2X engine health & status monitoring)
  - James et al. (Prognostic assessment component of BEAM)
  - Gorinevsky, Mah & Timuçin (Early detection of SRM failures via Kalman filter residual monitoring)
- No provision for false alarm mitigation !



- Pending paper:
  - Rodney A. Martin. Approximations of optimal alarm systems for anomaly detection. *IEEE Transactions on Information Theory (preprint under major revision for 2nd round submission)*, 2007.
  - Revised title will be "A State-Space Approach to Approximating Optimal Level-Crossing Predictions for Linear Gaussian Processes."

Motivation

- "Anomaly Detection" Problem
  - Some recently developed algorithms use thresholds not based on physical limits
  - Functional distinction must be made between design and failure thresholds

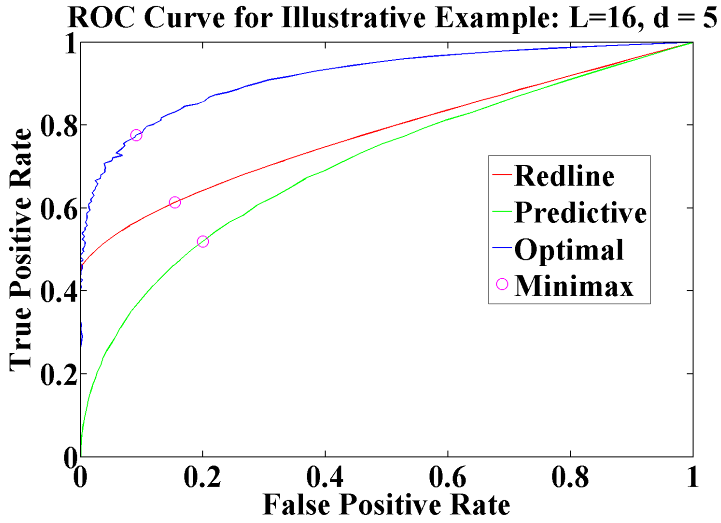


- Often resistance to maturing more advanced data-driven techniques, with some notable exceptions
- With optimal alarm, these concerns are addressing by incorporating legacy techniques and potentially more advanced data-driven techniques

Introduction to Optimal Alarm Systems

- Optimal Alarm Systems
  - Based upon level-crossing event defined over a fixed prediction horizon
  - Level-crossing event can be constructed as logical expression
  - Incorporates predicted future process values
  - Design alarm system that elicits fewest false alarms for fixed detection probability (Neyman-Pearson Lemma) and trade-off with prediction horizon using area under the ROC curve.

- Novelty/Advantages/Disadvantages
  - ✓ Uses recursive Kalman filtering
  - ✓ Design criteria can be written as function of:
    - ✓ Model parameters
    - ✓ Level-crossing based failure threshold
  - ✓ Adds predictive capability, and precise definition of optimality for alarms
  - ✓ Independent alarm system design and critical event definition
  - ❖ Must introduce approximations
  - ❖ Currently applicable to stationary linear systems driven by Gaussian noise
  - ❖ Not applicable to multivariate time series or able to exploit cross-correlations



Information Sciences & Technology

Ames Research Center

Current ARMD/Aviation Safety IVHM Impact

- Work considered for PART IVHM milestone 3.3.4 (Prognosis data mining: Forecasting technology that has the ability to predict at least 3 known anomalies in real or emulated data of large, fleet-wide heterogeneous data sources)
  - Advance the SOA in technologies related to predicting future anomalies
  - Baseline BEAM, which moves from "redlines" to prognostic assessment
  - Predictions provide confidence bounds on the exceedance of redlines, using theoretically and statistically sound approaches.
  - Cited to enjoy favorable false alarm rates, but no explicit theoretical allowance for false alarm mitigation

- Proposed Milestones
  - ❖ FY09: Theoretical investigation of approximations, comparative analysis using varied architectures and techniques
  - ❖ FY10: Improved approximations for alarm system design and improved machine learning techniques for model development
  - ❖ FY11: Extend optimal alarm systems to allow for introduction of particle filtering
  - ✓ FY12: PART Milestone Demonstration

Potential extensions to ARMD/Aviation Safety/IRAC

- Natural extension to conducting joint IVHM research w/IRAC into mitigation via modified aircraft operation (IVHM Milestone 3.4, Mitigation)

- Leverage previous research
  - Incorporate information at the time of alarm or other critical event into control actions via conditional expectation
  - Optimal control theory can be used via optimal reference modification
  - Optimal reference modification references:
    - Anders Hansson. Optimal modifications of reference signals for critical processes using alarm signals. In *Preprints IFAC 12th World Congress*, pages 615–618, Sydney, Australia, 1993.
    - Rodney A. Martin. *Optimal Prediction, Alarm, and Control in Buildings Using Thermal Sensation Complaints*. PhD thesis, University of California, Berkeley, 2004. (Ch. 6)

- New questions and ideas:
  - Control and convex optimization: study of convexity of alarm regions and feasibility for use as a constraint in this paradigm
  - Controllability and observability are sufficient conditions for construction of an optimal alarm system based upon an LG model…but what if we don't have an LG model ?
  - What happens when the system is nearing margins of stability due to degraded performance ?
  - Incorporate adapted dynamics into a new model to revise the failure propagation rate
  - Enforce response time of controller dynamics as constraints on prediction horizon