

# 基于 PHM 的电力中间件故障预警技术研究

张科<sup>1</sup>, 唐静<sup>2,3</sup>, 肖治华<sup>1</sup>, 廖荣涛<sup>1</sup>, 詹伟<sup>1</sup>, 李瑞轩<sup>2</sup>,  
向万红<sup>3</sup>, 朱闻远<sup>3</sup>, 宁立<sup>3</sup>

(1. 国网湖北省电力公司信息通信公司, 武汉 430077;

2. 华中科技大学 计算机科学与技术学院, 武汉 430074; 3. 远光软件股份有限公司, 广东 珠海 519085)

**摘要:** 在电力信息系统中, 中间件的运维工作需要从传统的事后排查提升为故障预判和智能预警, 面向电力中间件的故障预测与健康管理的 PHM 技术成为当前迫切需要研究的课题; 分析和利用 PHM 技术的数据处理流程, 突破性地将该技术应用于智能化运维管理平台的中间件集群管理; 以灰色状态下的马尔科夫预测模型为核心预测算法, 结合时间切片管理和动态置信阈值技术, 设计并实现了面向中间件集群管理的故障预警模型; 以湖北省电力公司构建的智能化运维管理平台上的实施应用为实例, 该方法实现对中间件故障准确预警, 并大大降低了虚警率; 实验结果证明了该方法的可行性和有效性。

**关键词:** 中间件; 故障预测与健康管理; 故障预警; 数据挖掘

## Research on Fault Early Warning Technology of Power Middleware Based on PHM

Zhang Ke<sup>1</sup>, Tang Jing<sup>2,3</sup>, Xiao Yehua<sup>1</sup>, Liao Rongtao<sup>1</sup>, Zhan Wei<sup>1</sup>, Li Ruixuan<sup>2</sup>,  
Xiang Wanhong<sup>3</sup>, Zhu Wenyuan<sup>3</sup>, Ning Li<sup>3</sup>

(1. Information and Communication Branch of Hubei Electric Power Company, Wuhan 430077, China;

2. School of Computer Science & Technology, Huazhong University of Science and Technology, Wuhan 430074, China;

3. YGsoft Inc., Zhuhai 519085, China)

**Abstract:** In power information system, the maintenance of middleware should be arranged from the traditional post inspection to the fault pre-diagnosis and intelligent early-warning. Fault prediction and health management (PHM) technology based on power oriented middleware has become an urgent research topic. Taking the middleware cluster as the research object, according to the data processing flow of PHM technology, researching on Grey-Markov forecasting, combining with time slice management and dynamic confidence threshold, this paper designs and implements a fault early warning model for the middleware cluster management. Taking the application of operations management platform in Hubei electric power company as an example, The method realizes the accurate warning of the middleware fault, and reduces the false alarm rate. The experimental results show that this method is feasible and effective.

**keywords:** middleware; prognostics and health management; monitoring and early warning; data mining

## 0 引言

目前国网公司统一推广的信息系统架构都是基于应用中间件连接底层数据库和上层应用, 中间件的运行稳定与否直接关系到整个信息系统的运行稳定, 对于信息系统中间件这一关键构建监管就显得尤为重要<sup>[1]</sup>。针对该问题, 目前主要使用人工的方式定期检查上述环境的工作状态, 力求及早发现故障及早解决。但是人工检查受制于时间、经验等因素, 检查时效性差, 准确度低, 更重要的是无法实现对故障的预测。面对着分散的信息资源无法直接有效地进行监管和隐患排查, 在发生问题时, 无法快速地准确定位问题症结所在, 进而影响公司信息系统运行考核指标。因此, 电力信息系统运维工作中实现信息

系统运维智能和故障预警和定位功能, 将是信息运维管理提升的核心内容。

PHM 技术兴起于设备和复杂装备维护领域<sup>[2-3]</sup>。在软件领域, 软件的故障预测也是软件工程的核心内容之一<sup>[4]</sup>。李楠等<sup>[5]</sup>提出将该项技术应用于软件领域, 利用软件运行的关键信息, 借助各种智能数据分析方法来健康状态, 在故障发生前进行预警, 并且进行了实践验证可行性。同时, 预测是 PHM 的关键技术, 其优劣对预警管理有着至关重要的作用<sup>[6]</sup>。基于灰色模型的故障预测技术已有较多的研究和实际验证<sup>[7]</sup>, 何厚伯等人研究了马尔可夫过程对有限状态转移的评估, 并验证了其可行性<sup>[8]</sup>。因此, 基于马尔可夫状态推演的灰色改进型预测是否有效, 将是本文验证的重点。

针对以上问题, 提出一种面向中间件的故障预测与健康管理的模型, 将当前电力信息系统中中间件的运维工作, 从传统的事后排查转变为故障预判和智能预警。依据故障预测与健康管理的 PHM 技术, 以灰色-马尔科夫为核心预测算法, 结合时间切片管理和动态置信阈值技术, 设计并实现了面向中间件业务的故障监控和预警系统流程。以湖北省电力公司构建的智

收稿日期: 2016-01-14; 修回日期: 2016-03-07。

基金项目: 国家电网公司湖北省电力公司科技项目 (52153314005W)。

作者简介: 张科 (1983-), 男, 湖北钟祥人, 博士研究生, 高级工程师, 主要从事信息系统架构设计、运行维护工作方向的研究。

能化运维管理平台上的应用为实例, 实践证明了该方案的可行性和有效性。

## 1 系统结构及原理

故障预测与健康理念类似于人体的健康管理, 通过定时体检, 了解人体的运行状态, 建立信息记录, 观察是否有疾病发生的征兆, 提早预防和治疗, 降低发病率、致残率和死亡率。面对中间件这样的软件系统也需要有类似的机制, 不但需要定时收集各个中间件的运行状况, 更应该通过长期的跟踪分析运行规律, 通过对状态和故障的数据分析, 了解并推测出中间件产品未来的状态和风险。

故障预测与健康理念 (PHM) 技术是一种全面的故障检测、隔离和预测技术, 它将状态监测、综合诊断、状态评估等技术进行整合, 着眼于数据的特征提取、数据融合、预测以提供性能预测评价和保障决策。我们通过如图 1 所示的 PHM 技术流程, 来实现对中间件进行故障预测和健康管理的。图中的虚框部分, 状态特征和状态分类步骤是采用专家经验的离线分析计算得到。

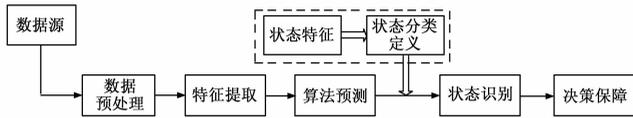


图 1 面向中间件的 PHM 流程

## 2 PHM 关键技术和实现过程

### 2.1 关键技术

一个完整的 PHM 体系应该包含数据采集, 数据预处理, 特征提取、数据预测、状态识别和故障评估决策几个步骤, 基于本文研究的面向中间件业务的故障监控和预警系统, 下面对关键技术和实施设计进行详细说明。

#### 2.2.1 数据采集和预处理

项目的目标是对中间件进行监测, 数据采集包括: 1) 中间件集群监测。分为 Cluster 数据监测、Nodestatus 数据监测、JDBC 数据监测、Heap 数据监测、Thread 数据监测和 Runtime 数据监测, 共六大类的性能指标监测; 2) 中间件节点指标监测。主要针对中间件集群中的各个 Server 数据监控和预警。通过 JMX 相关协议, 利用无插件方式抓取了中间件配置参数和实时运行状态的性能指标。针对异构数据, 我们进行格式化操作, 定义每项目标的统一数据结构, 原始数据序列为  $X = \{x_1, x_2, \dots, x_m, x_{m+1}, \dots, x_n\}$ , 其中  $\{x_1, x_2, \dots, x_m\}$  为训练数据,  $\{x_{m+1}, \dots, x_n\}$  为对比分析数据。以满足异构中间件指标数据的统一处理。

在数据预处理方面, 由于受中间件工作环境变化及系统工作状态的影响, 使监控数据中存在野值、随机误差等一些离散数据点, 这些离散的零点和奇异点会对预测结果产生影响。因此, 预处理可以通过滤波和插值等处理方式, 对其中的离散点数据进行处理, 提高后期预测判断的精准度。

预处理除了离散点处理, 还做了有关特征分析的预处理操作。对离散点预处理后的数据基于时间隙进行数据切片, 然后带时间戳进行特征提取。配合专家系统评判状态的多属性选择, 分析切片数据的多种统计特征属性。

#### 2.2.2 灰色状态下的马尔科夫预测模型

灰色预测算法对于短期预测的效果是普遍认同的, 但是由于传统的灰色 GM (1, 1) 预测算法是基于单变量的一阶微分方程的模型, 所以对较为线性的数据预测结果比较理想但是对于波动的数据序列, 预测结果就不不太令人满意。中间件监控数据由于设备的使用频率、使用时段等各种因素的影响, 数据波动性可能非常的大, 如果只采用灰色预测算法得出的结果进行分析, 可能对整个监控预警系统的误差造成较大的影响。在本文章将 GM 模型与马尔科夫模型结合即可得到灰色马尔科夫模型, 其中灰色可以给出时间序列的大体趋势, 数据序列围绕其进行波动, 对于波动产生的变化率可以引入状态转移矩阵, 引入马尔科夫过程进行进校正完成预测。

首先进行灰色建模预测, 设原始数据序列为  $X = \{x_1, x_2, \dots, x_m, x_{m+1}, \dots, x_n\}$ , 其中  $\{x_1, x_2, \dots, x_m\}$  为预测基础序列,  $\{x_{m+1}, \dots, x_n\}$  为对比序列。利用灰色算法对基础序列进行滚动预测, 获得预测结果序列  $\hat{X} = \{\hat{x}_{m+1}, \hat{x}_{m+2}, \dots, \hat{x}_n, \hat{x}_{n+1}\}$ 。得到相对波动变化率:

$$E_i = \frac{x_i - \hat{x}_i}{x_i}, i = m+1, \dots, n \quad (1)$$

将整个空间状态  $H$  根据变化率不同划分为  $k$  个等分的状态空间  $H = \{H_1, H_2, \dots, H_k\}$ ,  $H_i$  状态代表的变化率空间为  $[E_{\min} + (i-1) * (E_{\max} - E_{\min})/k, E_{\min} + i * (E_{\max} - E_{\min})/k, i = 1, 2, \dots, k$ 。那么当确定下一时刻为可预测状态时, 统计整理出以下矩阵:

$$\begin{bmatrix} M_{11} & M_{12} & \dots & M_{1m} & M_1 \\ M_{21} & M_{22} & \dots & M_{2m} & M_2 \\ M & M & M & M & M \\ M_{m1} & M_{m2} & \dots & M_{mm} & M_m \end{bmatrix} \quad (2)$$

其中:  $M_{ij}$  ( $i, j = 1, 2, \dots, m$ ) 表示在时间序列中前一时刻状态为  $H_i$ , 后一时刻为  $H_j$  出现的次数。由这个次数的统计矩阵就可以得到一次状态转移概率矩阵为:

$$P(1) = \begin{bmatrix} P_{11} & P_{12} & \dots & P_{1m} \\ P_{21} & P_{22} & \dots & P_{2m} \\ M & M & M & M \\ P_{m1} & P_{m2} & \dots & P_{mm} \end{bmatrix} \quad (3)$$

其中:  $P_{ij}(1) = M_{ij}/M_i, i = 1, 2, \dots, m$ , 若  $M_i = 0$ , 则  $P_{ij}(1) = 0, j = 1, 2, \dots, m$ 。

状态转移概率矩阵描述的是系统各状态转移的全部统计规律, 设在预测时刻  $n$  的状态为  $H_n$ , 则找到对应  $P(1)$  矩阵第  $n$  行中,  $\max(P_{nj}) = P_{nl}$ 。然后通过灰色预测结果和  $l$  值, 利用下面公式计算出最终预测值:

$$\hat{X}' = \frac{1}{2} * \hat{X} * [E_{\min} + (l+1) * (E_{\max} - E_{\min})/k + E_{\min} + l * (E_{\max} - E_{\min})/k] \quad (4)$$

### 2.2 面向中间件的 PHM 实现过程

针对中间件监控平台的实际情况, 本文设计和构建如图 2 所示的预测和故障预警流程。

图 2 中, 中间件预测与故障预警的步骤可总结如下:

- 1) 数据准备。首先需要建立一个通用的中间件参数模型, 采用无插件方式抓取了中间件配置参数和实时运行状态的性能指标。
- 2) 数据预处理和特征提取。基于对中间件运行业务逻辑理解, 在不破坏原始数据真实的基础上, 采用合适的方法解决

数据的冗余、不完整和不一致性问题。

3) 灰色马尔科夫预测。利用灰色算法对基础序列进行滚动预测, 获得预测结果序列  $\hat{X} = \{\hat{x}_{m+1}, \hat{x}_{m+2}, \dots, \hat{x}_n, \hat{x}_{n+1}\}$ 。然后, 依据相对波动变化率不同, 对整个空间状态  $H$  进行划分, 划分为  $k$  个等分的状态空间  $H = \{H_1, H_2, \dots, H_k\}$ 。最后, 计算状态空间的状态转移概率矩阵, 依据该矩阵确定序列未来的波动走向, 校正灰色预测结果  $\hat{X}'$ , 得出最后预测结果并计算绝对误差  $\epsilon = |\hat{X}' - X|$ 。

4) 计算置信区间。依据预测结果和绝对误差计算出对应时间切片区间内的置信阈值  $CI = \hat{X}'_{mean} + \alpha\epsilon_{mean}$ 。其中,  $\alpha$  为误差调整因子, 初始为数据分析结果设置, 后期将根据预警评估进行调整。

5) 预警。判断当前预测结果  $\hat{X}' > CI_{mean}$ , 如果是则预警;

6) 预警结果评估和反馈。发生预警后, 依据系统真实情况或者人工经验进行评估, 如果为虚警, 将结果返回步骤四, 重新计算  $\alpha$  因子。

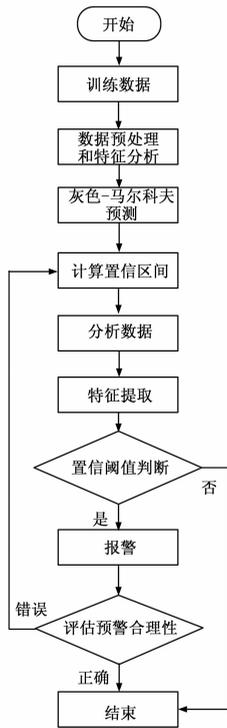


图 2 中间件预测与故障预警流程图

### 3 实例及结果分析

按照第二部分介绍的流程, 以湖北省电力公司构建的智能化运维管理平台中 heap 堆栈指标的监控和预警为例, 在该平台上建立基于中间件的智能监控和预警功能, 实现了故障预测和健康管理技术的应用, 以下介绍模型的实施效果并进行分析。

基于生产运行环境, 抓取基础 heap 堆栈指标的 3 个月实际数据, 数据抓取的频率为 1 分钟。基于数据安全考虑, 本文截取其中 300 个 heap 堆栈数据作为基础样例数据。经过无量纲处理后, 前 200 为训练数据, 后 100 为测试数据, 测试样例如图 3 (a) 所示。实际生产环境中, 在 3 点时对中间件进行压力测试, 是需要进行预警位置。图中的基础数据存在一定的信息缺失和奇异点, 针对基础数据的粗大误差和奇异点进行滤波, 然后针对数据中的零值和缺失值, 采用分段线性插值进行

处理, 弥补信息产生丢失, 结果如图 3 (b) 所示。

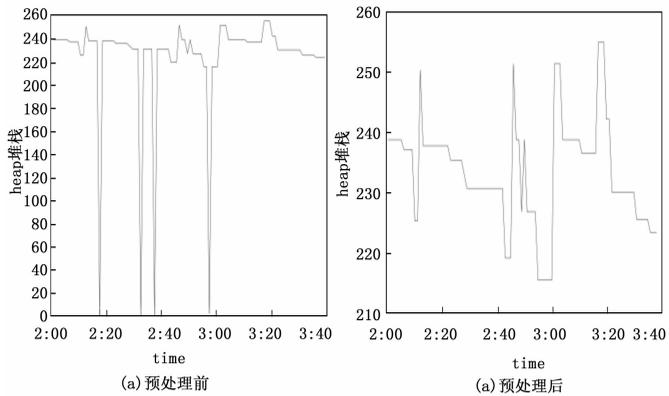


图 3 预处理前后的 heap 堆栈数据

预处理后, 采用灰色马尔科夫进行训练数据的计算, 预测出测试数据的预测值 (图 4 中点线所示)。同时, 根据每个预测点之前十个点的预测和真实值结果, 依据置信阈值公式计算得到置信阈值 (图 4 中虚线所示)。

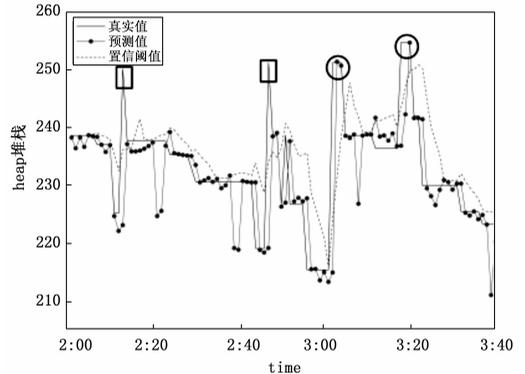


图 4 灰色马尔科夫预警

从图 4 中可以看出, 如果采用传统的静态阈值法, 将会在图中的方块标注位置出现虚警, 不易分辨。采用新的基于灰色马尔科夫预测、时间切片管理和动态阈值预警方法, 在虚警位置, 预测值并没有超过置信阈值, 可以准确避免虚警, 这个与实际情况相符。而在图中圆圈标注的位置, 预测值则明显超出了置信阈值, 系统实现了压力测试的准确预警。从结果可以看出, 该故障预测和健康管理方法可以对中间件实现准确预警, 并大大降低虚警率。

### 4 结论

在中间件监控系统中采用故障预测与健康管理技术, 是通过状态监测获取设备中间件状态和故障信息, 预测其故障发展趋势, 提高监管和排查的精准性。本文研究了基于灰色-马尔可夫算法的中间件状态预警模型, 运用马尔可夫模型方法, 将部件系统退化过程描述为有限状态转移过程, 建立了基于马尔可夫的健康状态评估模型, 利用时间切片管理和动态置信阈值技术实现对中间件监控和故障预警的功能。最后进行了案例分析, 结果验证了模型的可行性和可靠性。

#### 参考文献:

[1] 吴一鸣, 田永涛. 南宁供电局 J2EE 应用服务中间件监控系统的设

计与实现 [J]. 电信科学, 2013, 29 (11): 131-135.

- [2] 王 锐, 王 洁, 冯 刚, 等. 复杂装备故障预测与健康管理体系结构研究 [J]. 计算机测量与控制, 2012, 20 (7): 1740-1743.
- [3] 刘 晶, 季海鹏, 朱清香, 等. 含有隐性故障的复杂设备故障预警方法 [J]. 计算机测量与控制, 2014, 22 (4): 1030-1032.
- [4] Bonissone P, Iyer N. Knowledge and time: a framework for soft computing applications in PHM [J]. Uncertainty and intelligent information systems, World Scientific, 2007.

- [5] 李 楠. 故障预测与健康管理 (PHM) 在 ERP 系统中的应用研究 [D]. 合肥: 中国科学院大学 (工程管理与信息技术学院), 2013.
- [6] 单 珊, 冯玉光, 奚文俊. PHM 中预测性能评价方法的发展与展望 [J]. 计算机测量与控制, 2015, 23 (12): 3909-3912.
- [7] 黄大荣, 黄丽芬. 灰色系统理论在故障预测中的应用现状及其发展趋势 [J]. 火炮发射与控制学报, 2009 (3): 88-92.
- [8] 何厚伯, 赵建民, 许长安, 等. 基于马尔可夫过程的健康状态评估模型 [J]. 计算机与数字工程, 2011, 39 (7): 63-66.

(上接第 64 页)

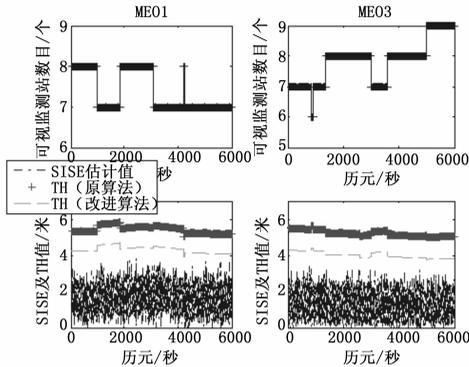


图 6 无故障模式下系统告警性能对比

模式下, 算法改进前后的实时 SISE 与完好性阈值 TH 的仿真结果如图 7、图 8 所示。其中, 两类故障的仿真时长均为 1 201 秒至 3 000 秒。

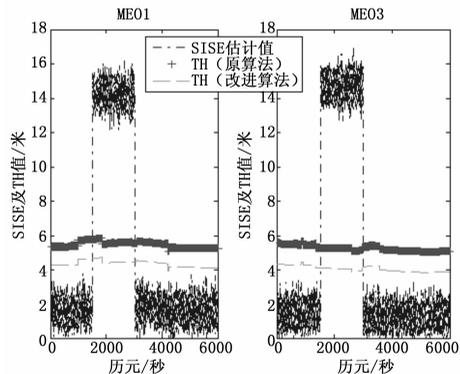


图 7 阶跃类故障模式下系统告警性能对比

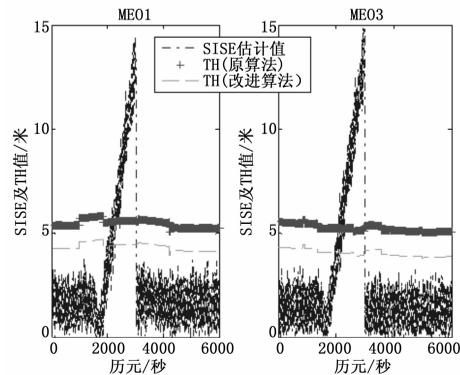


图 8 斜坡类故障模式下系统告警性能对比

积而增大, 仅当故障误差达到某一门限值, 实时 SISE 估计值超过阈值, 完好性标志 IF 置“1”, 系统告警。由于阶跃类故障的实时 SISE 值变化为瞬时值, 远短于 1 s, 两种算法均能及时告警。而对于斜坡类故障, 实时 SISE 值不断增大, 当 SISE 值超过 TH 值则系统告警。因此改进后算法能够更加及时告警。

### 4 结束语

文中系统地分析空间信号完好性监测系统的监测机理和算法模型, 针对其中两个关键技术点的不足之处, 提出优化方案: 改进临界圆法估计最坏用户位置 WUL; 基于非零均值 SISE 的 SISA 算法优化原算法过分包络问题。最后仿真分析无故障、阶跃类故障、斜坡类故障三种模式下, 算法改进前后系统的完好性告警性能。结果表明, 改进后的算法在保证 TH 对实时 SISE 的包络能力条件下, 将包络紧度由 3.89 米改善至 2.62 米; 且告警更加及时。

### 参考文献:

- [1] Grover Brown R. Global Positioning System Theory and Application [M]. American Institute of Aeronautics and Astronautics, 1996, 143-164.
- [2] GPS Interface Specification IS-GPS-200F [EB/OL]. <http://www.navcen.uscg.gov>.
- [3] Veit Oehler. The Galileo Integrity Concept [A]. Proceedings of the 17th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS 2004) [C]. 2004; 604-615.
- [4] Helmut Blomenhofer, Walter Ehret. Performance Analysis of GNSS Global and Regional Integrity Concepts [A]. Proceedings of the 16th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS 2003) [C]. 2003; 991-1001.
- [5] Shaojun Feng, Washington. An Efficient Worst User Location Algorithm for the Generation of the Galileo Integrity Flag [J]. Journal of Navigation, 2006, 59 (03): 381-394.
- [6] 谷守周, 秘金钟. Galileo 完备性系统中 WUL 算法研究 [J]. 测绘科学, 2011, 36 (2): 17-19.
- [7] 秘金钟. GNSS 完备性监测方法、技术与应用 [D]. 武汉: 武汉大学, 2010.
- [8] 张雪辉, 贝 超, 王 芳. 最差用户位置算法研究与仿真分析 [J]. 航天控制, 2010, 28 (2): 29-35.
- [9] Fan M J, Zhou J H. Algorithm of Analysis for the Improved Integrity Index SISMA of GALILEO System [A]. China Satellite Navigation Conference (CSNC) 2012 Proceedings [C]. 2012; 57-65.
- [10] Heng L, Gao G X, Walter T, et al. Statistical Characterization of GPS Signal-in-Space Errors [A]. Proceedings of the 2011 International Technical Meeting of the Institute of Navigation [C]. 2011; 312-319.

由上图 7 和图 8 可看出, 实时 SISE 估计值随故障误差累