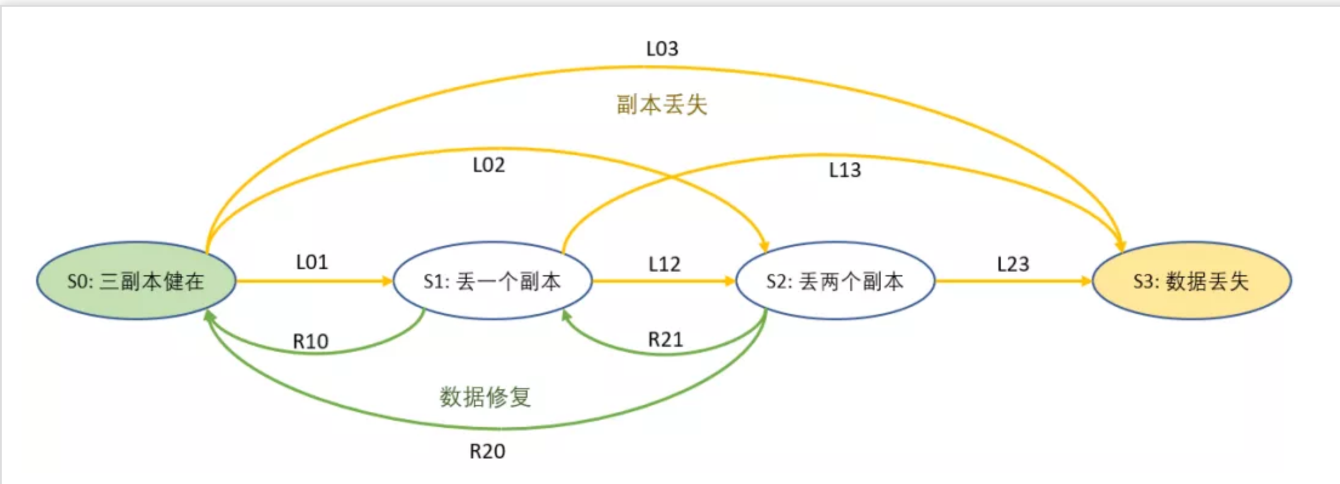
马尔可夫链的平均故障时间

Original Accela推箱子 Accela推箱子 2019-12-15

存储系统中，马尔可夫链（Markov Chain）常用于数据对象的生命周期建模，平均故障时间（MTTF）可评估存储的可靠性。在数据丢失的上下文中，也称作平均生存时间（Mean Time to Survival）。

例如三副本数据，随着更多服务器故障，逐步丢失全部副本；同时，健康服务器也试图重建丢失的副本。平均故障时间告诉我们期望一份数据多久不丢失（不进入S3状态）。



如何计算马尔可夫链的平均故障时间呢？感觉网上资料不多，讲明原理的少，所以写篇博文记录下:-D

计算方法

本文首先列出平均故障时间的计算方法，后面详细讲为什么。公式如下：

$$MTTF = S_0(I - M_s)^{-1}L_{k,0} \qquad (7)$$

模型定义

如何定义“故障”？它对应马尔科夫链的吸收状态（Absorbing State）：吸收状态不向自己之外的任何状态转移，例如数据丢失所有副本。

* 系统处于非吸收状态时表示存活，处于吸收状态时表示已经死亡，即“故障”。平均故障时间，或曰平均生存时间，指我们预期系统能在非吸收状态停留多久。

然后，用符号将马尔科夫链表示出来：

（1）式表示，单位时间dt内，从状态S_i转移到S_j的概率是p_{ij}。（“单位时间dt”说明p_{ij}不是概率，而是概率关于时间变化的导数。）

（2）式是马尔科夫链的状态转移矩阵。一共有n个状态结点。M[row=i][column=j]表示从状态结点i转移到j的概率。

（3）式是时刻t时，系统处于各个状态结点0...n的概率，概率之和恒为1。S_0是起始状态。

（4）式定义状态如何递进。（这个朴素的公式和特征向量、线性变换都能挂上关系）

$$P(S_i \to S_j) = p_{ij} * dt \qquad (1)$$

$$M = \begin{bmatrix} p_{0,0} & p_{0,1} & \cdots & p_{0,n-2} & p_{0,n-1} \\ p_{1,0} & p_{1,1} & & p_{1,n-2} & p_{1,n-1} \\ \vdots & & \ddots & \vdots & \\ p_{n-2,0} & p_{n-2,1} & & p_{n-2,n-2} & p_{n-2,n-1} \\ p_{n-1,0} & p_{n-1,1} & \cdots & p_{n-1,n-2} & p_{n-1,n-1} \end{bmatrix} * dt \qquad (2)$$

$$S_t = [S_{0,t} \quad S_{1,t} \quad \cdots \quad S_{n-2,t} \quad S_{n-1,t}] \qquad (3)$$

$$S_{t+dt} = S_t * M \qquad (4)$$

我们可以重新排列M的列，使得吸收状态总是位于M的最右边。定义状态结点k..n是吸收状态，M重写为如下:

（5）吸收状态的转移矩阵是单位矩阵，其左边总是零。M被分块，M_S是非吸收状态的转移矩阵。

$$M=\begin{array}{cc|l} \left[\begin{array}{cc}\left[\;M_S\;\right] & \left[\;*\;\right]\\ \left[\;0\;\right] & \left[\;I\;\right]\end{array}\right] & & \begin{array}{l}\}\,k\\ \}\,n\text{-}k\end{array}\\ \begin{array}{cc}\quad k\quad & \quad n\text{-}k\quad\end{array} & & \end{array} \qquad (5)$$

为了方便表示，我们引入一个计算工具$L_{a,b}$：将矩阵与它相乘，可以求和从0到a-1的列。

$$L_{a,b}=\begin{bmatrix}1\\ \vdots\\ 1\\ 0\\ \vdots\\ 0\end{bmatrix}\begin{array}{l}\left.\begin{array}{l}\\ \\ \\ \end{array}\right\}a\\ \left.\begin{array}{l}\\ \\ \\ \end{array}\right\}b\end{array} \qquad (6)$$

最后，我们可以给出本文开头的平均故障时间的计算公式（7）:

$$MTTF = S_0(I - M_s)^{-1}L_{k,0} \qquad (7)$$

注意MTTF的单位是“单位时间dt”。

* 例如，我们使用“每小时”某服务器故障概率为X，则最终计算出MTTF的单位是小时。

为什么这样计算

良好的定义能简化问题。首先我们对通用的故障问题定义如下

（8）在时刻t，系统存活的概率是P(t)。t之前系统一定没有故障，t之后系统可能故障也可能不。

（9）在时刻t邻域，系统有多大可能故障？单位时间dt内的故障概率被定义为h(t)。

（10）平均故障时间（MTTF）是存活时间长度的期望。它是求和tP(刚好在时刻t故障)。

$$P(t) := Probability\ of\ surivival\ at\ time\ t \qquad (8)$$

$$h(t) := \frac{Probability\ of\ dead\ at\ time\ t}{dt} \qquad (9)$$

$$MTTF := \int_{t=0}^{+\infty} tP(t)h(t)dt \qquad (10)$$

注意到h(t)和P(t)是有对应关系的，见（11）式。通过它，我们可以简化MTTF:

（12）MTTF可由存活概率的积分求得。tP(t)在t无穷大时必须为零，否则期望的存活时间是无限的。

$$P(t+dt)=P(t)*(1-h(t)dt)$$
$$\Rightarrow dP(t)=-P(t)h(t)dt \qquad (11)$$

$$MTTF=\int_{t=0}^{+\infty} tP(t)h(t)dt=-\int_{t=0}^{+\infty} tdP(t)$$
$$=\int_{t=0}^{+\infty} P(t)dt-[tP(t)]_0^{+\infty}$$
$$=\int_{t=0}^{+\infty} P(t)dt \qquad (12)$$

最简单的情况是什么呢？我们假设h(t)是定值，则P(t)符合指数分布，MTTF为1/h。指数分布是故障概率计算最常用的。

但马尔可夫链的h(t)不是常量。我们按照公式计算：

（13）时刻t的存活概率可以由状态转移矩阵累乘得到。矩阵、状态转移，还是需要引入离散。

（14）MTTF可以用离散形式计算，和积分形式是对应的，也能推导出（12）的简化形式。

$$m := \frac{t}{dt}$$

$$\mathrm{P}(t) = S_0 M^m L_{k,n-k} \tag{13}$$

$$\begin{aligned} MTTF &= \sum_{m=0}^{+\infty} tP(t)h(t)dt = \sum_{m=0}^{+\infty} t(P(t) - P(t+dt)) \\ &= \sum_{m=0}^{+\infty} P(t)dt - [tP(t)]_0^{+\infty} = dt \sum_{m=0}^{+\infty} P(t) \end{aligned} \tag{14}$$

有一些通用的公式可以帮助进一步化简（14）:

（15）利用矩阵M的吸收状态的分块，我们可以把M的累乘化简为M_s的累乘

（16）矩阵累乘的叠加可以转换为一个逆矩阵。通过等式两边乘以逆矩阵来证明。

$$M^m L_{k,n-k} = {M_s}^m L_{k,0} \tag{15}$$

$$I + M_s + {M_s}^2 + {M_s}^3 + \cdots = (I - M_s)^{-1} \tag{16}$$

运用（15）和（16），逐步化简（14），最终我们得到本文开头的公式（7）。注意（7）中MTTF的单位是“单位时间dt”。

$$MTTF = dt \sum_{m=0}^{+\infty} S_0 M^m L_{k,n-k} = dt \sum_{m=0}^{+\infty} S_0 {M_s}^m L_{k,0}$$
$$= S_0 \left(\sum_{m=0}^{+\infty} {M_s}^m \right) L_{k,0} * dt = S_0 (I - M_s)^{-1} L_{k,0} * dt \tag{17}$$

Cheers~ :-D

相关资料

[Calculation of MTTF values with Markov Models for SafetyInstrumented Systems](http://www.wseas.us/e-library/conferences/2007venice/papers/570-618.pdf)

[Wikipedia: Stochastic matrix](https://en.wikipedia.org/wiki/Stochastic_matrix)