

基于时效量化的 Web 服务评价模型

印桂生 崔晓晖 马志强 董宇欣

(哈尔滨工程大学计算机科学与技术学院 哈尔滨 150001)

摘 要 现有的服务评价方法缺乏时效性,而应用于其他系统的时效量化方法难以体现服务评价时效在衰退上的差异特征。由此,建立了基于时效量化的 Web 服务评价模型。模型以评价时效的遗忘特征为切入点,分析评价时效与人类记忆的关系,揭示评价时效遗忘速率的更新规律,构建符合时效衰退差异的时效量化方法。同时,根据量化结果,优化评价窗口长度,从而解决人工设定窗口长度的不合理性。实验表明,模型在满足时效衰退差异性的基础上,确保了较低的评价误差和计算代价,并合理拟合用户评价,具有一定理论研究和实际应用价值。

关键词 Web 服务,服务评价,时效量化,遗忘过程,评价窗口

中图分类号 TP311 文献标识码 A

Web Services Evaluation Model Based on Qualification of Time Utility

YIN Gui-sheng CUI Xiao-hui MA Zhi-qiang DONG Yu-xin

(College of Computer Science and Technology, Harbin Engineering University, Harbin 150001, China)

Abstract Current evaluation methods of Web services do not consider the time-utility, while qualification of time-utility in other systems are hard to satisfy the requirements of distinction of decay of time utility. As a result, Web service evaluation model based on qualification of time utility was proposed. Based on forgetting phenomenon of time utility, the model analyzes the relationship between time-utility and memory, unfolds regular of update of forgetting-ratio to time-utility, and designs method to qualify time utility concerned with the distinction of decay of time utility. Meanwhile, result of time utility is used to optimize calculation of length of evaluation windows, as to improve KNN in artificial setting of evaluation windows. Experiments shows that model meets the requirements of distinction of decay of time utility, achieves low evaluation error, reduces computational cost and fits users' evaluation, which is meaningful to theoretical research and practical use.

Keywords Web services, Service evaluation, Qualification of time utility, Forgetting process, Evaluation windows

1 引言

Web 服务是一种运行于 Web 上的自包含、自描述、模块化的应用程序,可以在网络中被描述、发布、查找以及调用^[1]。随着分布式软件技术的发展,适用于各领域的 Web 服务大量涌现。那么,如何发现用户所需的 Web 服务,是服务评价领域的研究重点。

现有方法主要从两方面获取 Web 服务运行时的综合评价:(1)侧重评价指标的研究,旨在通过多目标决策过程降低问题的复杂性,比较有代表的成果包括 WSES^[2] 和 Qos^[3-5] 等。(2)借鉴社会学中参与信任过程的角色关系建立模型^[6,7]。其基本过程是,当用户发出服务评价请求时,根据服务实体的历史行为,从失效率和成功率等可信度上衡量 Web 服务综合性能评价。

上述方法在计算综合评价时,通常采用类 KNN^[8] 等评价窗口手段分析数据。这种手段将发生在不同时刻的不同服务

评价等同对待,缺乏对其时效性的考虑。一种自然的想法是以时间作为效用因子,来衡量不同服务的时效性,据此建立历史信息时效演化关系^[9,10]。在一些研究领域,已存在能够满足时效衰退的数学模型,例如应用于推荐系统中的幂函数^[11]、应用于专家系统中的核函数^[12]及应用于概念漂移系统中的线性函数^[13]。

根据 Indre Zliobaite^[14] 研究,评价具有不确定性和差异性,这种差异性来源于人们对服务完成指定功能的一种信念。如果采用时效量化来拟合这种差异特征,则表现为不同服务时效衰退差异。现有的时效量化模型均采用同一过程描述时效衰退,无法满足该衰退差异的需求。由此,本文建立基于时效量化的 Web 服务评价模型(WSEM-QTU, Web Services Evaluation Model based on Qualification of Time Utility),将时效量化过程与人类遗忘特征相结合,设计出遗忘速率动态调整方法,并以此建立体现不同服务评价时效衰退差异的时效量化方法。同时,引入时效量化结果,优化 KNN 过程,计

到稿日期:2011-06-23 返修日期:2011-09-28 本文受国家自然科学基金项目(60973075),黑龙江省省自然科学基金项目(F200937),哈尔滨市科技局基金项目(RC2009XK010003)和哈尔滨工程大学中央高校基本科研业务费专项基金项目(HEUCF1015, HEUCF100605)资助。

印桂生(1964—),男,博士,教授,主要研究方向为服务发现和可信软件等,E-mail:yinguisheng@hrbeu.edu.cn;崔晓晖(1984—),男,博士生,主要研究方向为服务评价;马志强(1977—),男,讲师,主要研究方向为分布式系统;董宇欣(1974—),女,副教授,主要研究方向为网构软件。

算评价时的窗口长度,从而降低计算代价,并获得更加拟合用户预期的服务评价结果。

本文第2节分析评价时效与人类遗忘的关系,介绍WSEM-QTU的形式化表示;第3节设计时效量化方法,阐述应用时效优化服务评价窗口长度的方法;第4节进行仿真实验,分析本模型的评价效果;最后总结全文内容并指出进一步的研究方向。

2 基于时效量化的 Web 服务评价模型

德国心理学家 H. Ebbinghaus^[15]率先采取无意义音节作为材料,以重学法记录在间隔一段时间后再对相同材料重新学习所花费的时间,揭示记忆量随时间变化的情况,如图1所示。

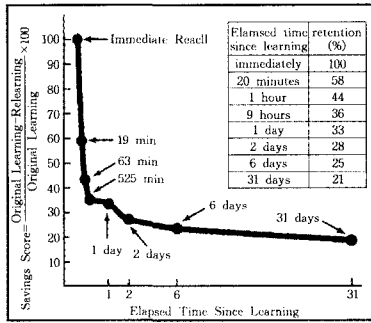


图1 遗忘曲线

服务评价时效与记忆的遗忘特征具有相似关系。首先,当用户使用他人感受作为选取服务的依据时,很自然会优先考虑近期的评价,从而忽略了已过时的评价。该过程表明,新的评价往往比旧的评价更具有参考性,即评价的时效遗忘特征。其次,某些被试材料在记住后很长时间都不会忘记,而有些材料即使记牢,仍然会很快忘记,即不同被试材料存在遗忘差异性。如果将被试材料看成服务评价时效,则这种差异性恰好体现出服务评价时效衰退差异。因此,本文参照记忆心理学中关于人类遗忘的研究成果,来描述 Web 服务评价时效随时间变化特征,具有一定实际性和参考性。

形式上,WSEM-QTU 可定义为 6 元组 $\langle U, W, R, S, \delta, N \rangle$ 。其中, $U = \{u_i\}$ 表示用户集合; $W = \{w_j\}$ 表示服务库中的 Web 服务集合; $R = \{r_{(i,j,t)}\}$ 表示时刻 t 用户 u_i 对服务 w_j 的评价反馈集合,可形式化为 $u_i \times w_j \times t \rightarrow r_{(i,j,t)}$; $S = \{s_j(t, k)\}$ 为服务评价时效量化函数集合, $s_j(t, k) \in (0, 1]$ 为指定服务 w_j 在时刻 t 的评价时效, k 用于控制该服务的时效衰退速率; δ 表示遗忘速率调整的参数。 N 表示模型在进行服务评价时,可供参考的最大评价窗口长度, $N \in \mathbb{Z}^+$ 。WSEM-QTU 根据服务评价时效量化结果,优化评价窗口长度,从而获取拟合用户期望服务评价的结果。

WSEM-QTU 的运行分为两个阶段:时效量化和服务评价计算。

时效量化:根据用户评价请求中的服务标识 j ,获取当前时刻 t 该服务 w_j 的遗忘速率 k ,计算 $s_j(t, k)$ 。其中, k 体现了不同服务在评价时效衰退的差异,是本文所需解决的核心内容,其具体求解方法将于第3节中介绍。

服务评价计算:WSEM-QTU 计算服务评价的具体过程如式(1)所示。

$$\begin{cases} r_{(i,j,t)}^n = \frac{\sum \{r_{(i,j,t')} \mid T \max(r_{(i,j,t')})\}}{n} \\ n = f(s_j(t, k), N) \end{cases} \quad (1)$$

式中, $r_{(i,j,t)}^n$ 表示在时刻 t 模型计算得到的服务评价价值; $\{r_{(i,j,t')} \mid T \max(r_{(i,j,t')})\}$ 用于获取服务 w_j 距离当前时刻 t 最近的 n 条用户评价记录; n 为模型采用时效 s_j 和系统参数 N 优化后的评价窗口长度; f 为优化函数,是本文所需解决的核心内容,其具体求解方法将于第4节介绍。

3 基于时效量化的 Web 服务评价模型

如果将时效随时间遗忘的现象看作是信息量不断流逝的过程,则可建立室来分析该过程。其中,室是存储评价时效的空间,它随时间依照遗忘速率 k 进行衰退,该过程如图2所示。

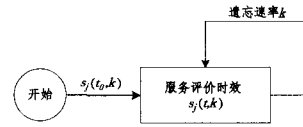


图2 时效衰退分析

图2表示了 w_j 的时效 s_j 随时间 t 不断衰退的动态过程。描述该过程的微分方程如式(2)所示。

$$ds_j(t, k)/dt = -ks_j(t, k) \quad (2)$$

根据遗忘特征, $s_j(t, k)|_{t=t_0} = 1$ 。代入并求解式(2), s_j 的原函数如式(3)所示。

$$s_j(t, k) = e^{-kt} \quad (3)$$

式(3)的图像符合 Ebbinghaus 实验结果,揭示了服务评价时效按照人类遗忘特征衰退的规律。因不同服务在时效衰退方面存在差异,故在式(3)的基础上,还需设计满足该特征的服务评价时效量化方法。

在记忆遗忘过程中,新的刺激(如重新复习已经学过的内容)会使已遗忘的记忆量重新回到完整状态。在后续的时间中,遗忘曲线以更加平缓的形式出现,即刺激后的阶段具有更低的遗忘速率,该过程称为记忆加强,如图3所示。

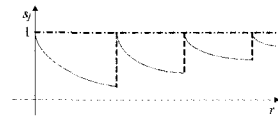


图3 记忆加强现象

记忆加强现象从实验上证明了不同的被试材料经过学习,会具有不同的遗忘特征,体现了时效的差异特征。根据该原理,在收到用户评价反馈时,WSEM-QTU 对式(3)中的遗忘速率进行更新,从而满足时效衰退差异的需求。

截取图3中相邻两个时效遗忘阶段,平移后一阶段的曲线,使之与前一阶段的曲线具有相同起点,以便分析遗忘速率的更新过程。该过程的示意图如图4所示。

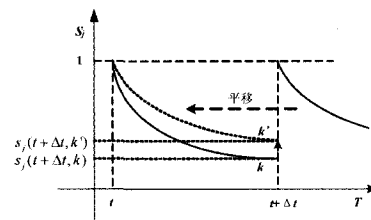


图4 遗忘速率更新过程

在图4中, t 时刻时效以 $s_j(t, k)$ 开始衰退。在 $t + \Delta t$ 时刻, 模型收到用户反馈, 原时效速率 k 被更新为 k' 。从 $t + \Delta t$ 时刻起, 时效以 $s_j(t + \Delta t, k')$ 进行衰退。若令 $\beta = s_j(t + \Delta t, k') - s_j(t + \Delta t, k)$, 表示曲线的更新程度, 则由式(3)得式(4)。

$$\beta = e^{-k'(t+\Delta t)} - e^{-k(t+\Delta t)} \quad (4)$$

化简式(4), 由 β 间接建立 k' 和 k 的递推关系, 如式(5)所示。

$$k' = \frac{\ln(\beta + e^{-k(t+\Delta t)})}{-(t+\Delta t)} \quad (5)$$

引入 WSEM-QTU 中参数 δ , 消去式(5)中 β , 得:

$$\beta = \frac{1 - e^{-k(t+\Delta t)}}{\delta} \quad (6)$$

在式(6)中, δ 的实际意义是每次更新程度占可更新程度上限的比重。其值越大, 则用户评价反馈对时效遗忘速率的调整程度越小。通常, δ 可以在整个服务评价系统中设定为一个大于1的正整数。将式(6)带入式(5), 得到 k' 的递推计算公式:

$$k' = \frac{\ln(1 + (\delta - 1)e^{-k(t+\Delta t)}) - \ln\delta}{-(t+\Delta t)} \quad (7)$$

在给定初始遗忘速率 k_0 的情况下, 任意时刻服务评价时效遗忘速率可通过式(7)的递推关系计算。若令时间序列 $t = \{t_0, t_1, \dots\}$ 表示服务 w_j 收到用户评价的反馈时刻, k_m 表示相邻时刻 t_m 和 t_n 间的遗忘速率, 则式(3)所示的时效量化函数可推广为:

$$s_j(t, k_m) = e^{-k_m(t-t_m)} \quad t \in [t_m, t_n] \quad (8)$$

式(8)即满足了服务评价时效衰退的需要, 又具有差异化时效衰退特征, 符合 WSEM-QTU 对时效量化的要求。

4 服务评价窗口优化

现有服务评价模型的数据分析方法, 通常采用人工设定评价窗口 n 的长度。若窗口太长, 则造成计算资源浪费; 若窗口太短, 则无法保证当前评价的准确度。本文根据时效量化结果, 使用函数 f 优化 n 的计算过程, 从而解决 KNN 在窗口长度设定方面的不合理性, 降低服务评价的计算代价。

为确保优化函数 f 在设计上的合理性, 需分析 n 与 $(s_j(t, k_m))'$ 的因果关系。

(1) n 与 Δt 的因果关系。根据式(8)函数递减特性, Δt 越大, 表明本次用户评价距离上一次的用户评价时间越长, 为保证本次评价计算的可靠性, 评价值的更新需要更多的历史记录作为参考。反之, Δt 越小, 则确定新的评价就需要越少的历史信息作为参考, 即 $n \propto \Delta t$ 。

(2) Δt 与 $(s_j(t, k_m))'$ 的因果关系。如果将 $(s_j(t, k_m))'$ 与 Y 轴截距看作是长度为 k_m 的一段线段, 则该线段上距离 $-k_m$ 越近的点, 表明其对应的 Δt 越小, 即 $\Delta t \propto (s_j(t, k_m))'$ 。由于 $n \propto \Delta t$, 则 $n \propto (s_j(t, k_m))'$ 。

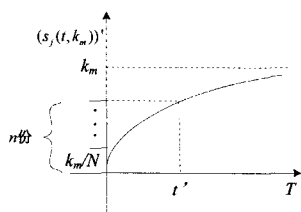


图5 n 计算示意图

根据以上推理, 只需构造满足 $n \propto (s_j(t, k_m))'$ 的等式, 即

可合理计算 n 。构造过程是: 首先将 $(s_j(t, k_m))'$ 的各点向上平移 k_m 个单位, 使得 $(s_j(t, k_m))' + k_m \in [0, k_m]$ 。其次, 将 $(s_j(t, k_m))' + k_m$ 平均分成 N 份, 每份长度为 k_m/N 个单位。根据 $n \propto (s_j(t, k_m))'$, 由连续函数的介值定理, 在 t' 时刻 WSEM-QTU 进行服务评价时, 恰好存在 n 份以 k_m/N 为单位的长度与之对应。该过程如图5所示, 上述数值关系如式(9)所示。

$$\begin{cases} (s_j(t', k_m))' + k_m = n \frac{k_m}{N} \\ (s_j(t', k_m))' = -k_m s_j(t', k_m) \end{cases} \quad (9)$$

由式(9), 在 t 时刻下, 优化函数 f 如式(10)所示。由于窗口大小必须是整数, 因此通过取整获取服务评价所需窗口长度。

$$n = f(s_j(t', k_m), N) = \lceil N(1 - s_j(t', k_m)) \rceil \quad (10)$$

式(10)在原有 KNN 的基础上, 充分考虑了时效量化对服务评价过程影响, 降低了评价过程的窗口代价。

5 仿真分析

5.1 数据准备及参数设定

随着终端技术的发展, 大量具有 Web 服务特征的应用涌现于 App Store^[16] 中, 如新闻应用、天气应用和订购应用等。本文采用这些应用的评价数据分析 WSEM-QTU 的评价效果。因 App Store 采用信用卡的身份认证方式, 且有效评价只产生于使用后的用户, 而大量应用需支付使用费用, 这些因素确保了评价数据的真实性和可靠性。

仿真过程使用绝对误差(MAE, Mean Absolute Error)分析仿真结果, 如式(11)所示, $r_{(i,j,t)}^m$ 表示由评价模型计算的结果, $r_{(i,j,t)}$ 表示标准评价。显然, MAE 的值越小, 说明评价效果越优秀。

$$MAE = \frac{\sum |r_{(i,j,t)}^m - r_{(i,j,t)}|}{M} \quad (11)$$

在仿真数据获取方面, 通过 Search API 抓取新闻类别下的 10000 条评价数据作为实验数据集。按照被评应用, 对数据进行预处理, 选择获取到评价最多的应用作为分析对象。根据实际预处理结果, 采用 $id=50$, 评价数据量为 583 条的数据集作为实验数据集。仿真中 WSEM-QTU 的各参数取值情况如表1所列。

表1 参数设置

参数名称	取值大小
N	5
k_0	1
δ	50

5.2 不同时效量化方法的评价效果分析

本节采用文献[11]提出的时效量化方法来构建与本文模型类似的服务评价模型 WSEM-E(k), 验证其与 WSEM-QTU 在实验数据集上的评价效果。两模型的差别在于, 前者在服务评价过程中直接采用无衰退差异的时效量化结果, 而后者采用差异化的时效量化结果。

为显示不同参数控制下 WSEM-E 的评价效果, 选取 $k=1$ 和 $k=0.5$ 分别进行实验, 验证其与 WSEM-QTU(our)在数据集上所得 $r_{(i,j,t)}^m$ 和用户标准评价 $r_{(i,j,t)}$ 的最小二乘拟合结果, 如图6所示。根据拟合后的实验结果, 不同参数作用下, WSEM-E(k) 计算结果的变化趋势相同, 其拟合曲线与标准评

价值的拟合曲线(standard)差距较大。对比 WSEM-QTU,其拟合曲线与标准拟合曲线较为接近。

图 7 分析了 WSEM-E 和 WSEM-QTU 在实验数据集上评价误差 MAE 的 6 次采样结果,前 5 次数据分段长度为 100,第 6 次为 83。分析 MAE 数值情况(见图 7),在不同参数下,WSEM-E 的 MAE 随评价次数增加发生波动,评价结果拟合用户标准评价的效果较差。对比 WSEM-QTU,在服务评价过程中,评价效果均优于 WSEM-E。583 次预测后,MAE=0.922,MAE 趋于稳定状态。

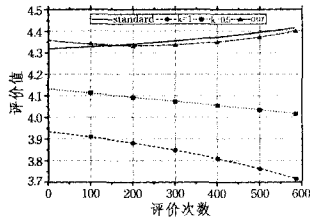


图 6 评价效果拟合分析

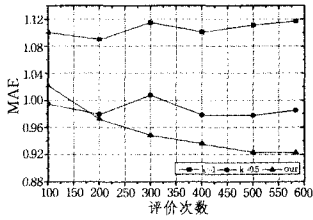


图 7 评价效果 MAE 分析

由上述分析,将现有时效量化方法直接引入服务评价过程中,并不能体现其时效的差异性,导致评价预测效果反而下降。而 WSEM-QTU 在满足服务评价时效差异性的同时,合理拟合用户标准评价。实验中,其具有较好的评价效果。

5.3 评价窗口优化效果分析

本节采取未经过时效优化的 WSEM-KNN 作为对比,分析其与 WSEM-QTU 在实验数据集上的评价效果及评价窗口代价。

在图 8 中,WSEM-KNN 每次选取与评价窗口长度相等的历史评价数据集作为当前服务评价计算参考,其 MAE 值在第 300 次左右评价后开始收敛,经过 583 次评价后,其保持在 0.970。对比 WSEM-QTU,由于在窗口计算中引入时效量化方法,其 MAE 在 500 次评价后开始收敛,表现为较慢的收敛速度。但在评价效果方面,仅在前 100 次微劣于 WSEM-KNN,其余各评价阶段均比 WSEM-KNN 具有更好的预测效果。

图 9 说明 WSEM-KNN 和 WSEM-QTU 在服务评价过程中所需要付出窗口代价的累加程度。窗口代价的实际意义是服务评价的计算代价。根据仿真结果,WSEM-KNN 的窗口代价随评价次数呈现线性上升,直线的斜率为窗口长度。经过 583 次评价后,WSEM-KNN 的窗口代价总量为 2915。比较而言,经过时效优化的 WSEM-QTU 所需要付出的窗口代价要小得多。在仿真数据集上,583 次评价付出的窗口代价仅为 750,比相同参数设定下的 WSEM-KNN 减少了 74.27% 计算代价。

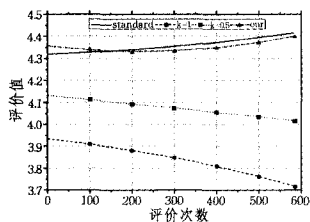


图 8 不同评价窗口策略的 MAE 分析

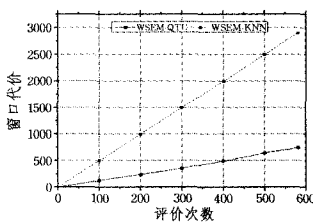


图 9 窗口代价分析

由上述仿真结果分析,与 WSEM-KNN 相比,WSEM-

QTU 可根据服务评价时效量化结果,合理地优化评价窗口长度,从而在减少评价计算代价的同时,获得更加拟合用户需求的评价结果。

结束语 针对服务评价过程中的时效问题,提出基于时效量化的服务评价模型。该模型能够根据遗忘速率更新方法,获取满足服务评价衰退差异特征的时效量化结果。并利用该结果优化评价窗口长度,从而在降低评价计算资源消耗的同时,获取更加拟合用户需求的服务评价。实际中,评价时效不仅具有简单衰退性,还具有其他复杂性质,如突发性记忆、群体记忆和联想记忆等。在下一阶段,拟从社会网络的角度出发,研究时效的社会性质,从而确定更加符合实际时效变化的服务评价模型。

参考文献

- [1] 孙萍,蒋昌俊. 利用服务聚类优化面向过程模型的语义 Web 服务发现[J]. 计算机学报,2008,31(8):1340-1353
- [2] 杨文军,李涓子,王克宏. 领域自适应的 Web 服务评价模型[J]. 计算机学报,2005,28(4):514-523
- [3] Kiran D, Lecturer M T. A greedy approach with criteria factors for QoS based Web service discovery[C]// Proceedings of the 2nd Bangalore Annual Compute Conference. NY: ACM, 2009:1-5
- [4] 张莹,黄厚宽,杨冬,等. 基于 Chord 的带有 QoS 的语义 Web 服务发现方法研究[J]. 电子与信息学报,2009,31(3):711-715
- [5] 周敏,张为群,林已杰,等. 一种基于扩展 Owl-S 本体的 Web 服务质量度量及评价方法的研究[J]. 计算机科学,2010,37(05):127-129,183
- [6] 邓忠军,王少杰,郑雪峰,等. 一种基于个体经验的多粒度信任模型[J]. 计算机科学,2010,37(04):91-94,105
- [7] 赵生慧,吴国新,陈桂林,等. 增强 Web 服务可信性的服务选择方法[J]. 通信学报,2011,32(1):37-45
- [8] Segaran T. Programming Collective Intelligence[M]. 南京:东南大学出版社,2008:167-196
- [9] 李小勇,桂小林. 动态信任预测的认知模型[J]. 软件学报,2008,21(1):163-176
- [10] 于洪,李转运. 基于遗忘曲线的协同过滤推荐算法[J]. 南京大学学报:自然科学,2010,46(5):520-527
- [11] Ding Yi, Li Xue. Time Weight Collaborative Filter [C]// Proceedings of the 14th ACM International Conference on Information and Knowledge Management (CIKM' 05). Bremen, Germany: ACM, 2005:485-492
- [12] Kukar M, Si L. Drifting Concepts as Hidden Factors in Clinical Studies[C]// Artificial intelligence in Medicine 9th Conference on Artificial Intelligence in Medicine in Europe(AIME 2003). Protaras, Cyprus: IEEE, 2003:355-364
- [13] 郑先荣,曹先彬. 线性逐步遗忘协同过滤算法的研究[J]. 计算机工程,2007,33(6):72-73,82
- [14] Zliobaite I. Learning under Concept Drift: an Overview [R]. USA: Faculty of Mathematics and Informatics, Cornell University, 2010
- [15] Ebbinghaus H. Memory: A Contribution to Experimental Psychology[J/OL]. <http://psy. ed. asu. edu/~classics/Ebbinghaus/index. htm>
- [16] Apple. App Store[OL]. <http://www. apple. com. cn/mac/app-store,2011-05-12>