

一种基于小世界网络中的 Web 服务智能路由算法^{*})

徐新卫 徐晓明 周 良 丁秋林

(南京航空航天大学计算机应用研究所 南京 210016)

摘 要 依据小世界网络具有特征路径短、聚类度大的特点和服务的聚集性,将提供 Web 服务的对等节点构造成为具有小世界属性的网络,在服务的查询、路由中可以提高效率和减少通信量。Web 服务路由凭借其虚名称和路由感知机制,在 SOAP 中方便地实现路由路径的指定和装载。本文借助蚁群算法,提出了在小世界网络模型中的基于 QoS 约束下的路由算法(QAA),通过实验仿真和算法分析表明 QAA 在对等的小世界网络中具有高效的路由性能。

关键词 Web 服务, Web 路由, 小世界网络, 蚁群算法

An Algorithm of Web Routing in Small-worlds Network

XU Xin-Wei XU Xiao-Ming ZHOU Liang DING Qiu-Lin

(Computer Application Institute, Nanjing University of Aeronautics and Astronautics, Nanjing 210016)

Abstract Because the small-world networks have a short length of characteristic path and evident attention of clustering, and Web service cluster together easily, the efficiency of communication is improved in Web services querying and routing on the network which has small-world features and is constructed by some peer to peer services that regard as vertex in graph. Web service depends on virtual name and routing aware mechanism to expediently implement routing appointed and loaded in SOAP. Based on basic ant colony algorithm, the QAA (QoS constrained Ant colony Algorithm) is presented. An experiment and analysis shown that QAA has high efficiency to Web service routing in the egalitarian small-world network.

Keywords Web service, Web routing, Small-world network, Ant colony algorithm

1 引言

Web 服务是由业务服务和业务组件组成的单元,是当今企业的下一代信息技术战略。该技术架构在 XML(eXtensible Markup Language)之上,并且可通过 SOAP 协议序列化成为 HTTP 的信息,跨越不同服务平台和组织机构^[1~3]。互联网的发展将使网络变得越来越丰富和密集,门户节点逐步消融,形成一个更加和谐的自然网络 Web 服务作为对等计算的对象是网络技术的发展趋势。对等计算是目前协调网络资源和分布式计算解决方案之一,使得服务具有亲和力、可操作性、高可用性,并增强用户体验,成为下一代 Web 服务的运行环境。研究中发现,Web 网页的内容、网页的链接以及用户在浏览 Web 页时产生的数据,具有小世界网络的性质。

“六度分离”(six degrees of separation)是小世界网络一个特征表现,指的是在网络中任意两个节点之间平均有六个节点将它们分开,表现网络在空间上的聚集性(Clustering)和分离度。1998 年,美国康乃尔大学 Watts & Strogatz(W&S)在“Nature”上发表了关于小世界理论的论文^[4],在国际上掀起了网络研究的热潮。几乎同时,法鲁托斯兄弟在观察 Internet 网、巴拉巴西研究小组在观察万维网的过程中发现:节点/网页对连接数是幂次律的关系,并且这一规律存在于各种网络模式^[5,6]。因为在 Web 服务中 SOAP 协议不处理显式的路由,所以 WS-Routing 和 WS-Referral 用于规定消息路径及其动态配置^[7]。在对等模式和小世界网络特征的影响下,定位服务和发现服务是至关重要的。

在 Web 服务中, QoS 是影响其路由路径选择的关键,也是 Web 服务实现的优化体现。本文的研究内容是以 Web 路由机制为出发点,构造具有小世界网络特性的网络模型的基础上,利用基于 QoS 约束的蚁群算法实现 Web 的智能路由。通过实验仿真,这种路由机制能很好地适应具有小世界性质的对等网络中。

2 小世界网络

小世界网络的拓扑结构介于全连接的规则网络和随机网络之间。在规则网络中体现随机性,主要有两种结构,即平等式(egalitarian)网络和无标度(free-scale)网络。平等式网络中的节点分布基本上是均匀的;无标度网络有极少数节点占用大部分的连接(按幂率),具有中枢节点和弱连接点。WWW 的网络就是无标度网络。根据 Web 服务具有相近性对服务进行聚类:服务将以相近的服务构成服务簇(services clustering),确定服务簇的关键节点(又称 hub 节点),服务簇的关键节点组成环路,随机引入关键节点的捷径,使得构造除的服务网络具有小世界网络属性^[8,9]。在小世界网络中不考虑网络拓扑结构而节点间有很小的平均距离;并且有较大的相近系数,使得网络节点之间通信的信息量将会减少,并且路由过程中涉及的中间节点数量也不多,适合 Web 路由的实现。对每一个服务簇的处理方式的不同,或 QoS 的要求不同,采用蚁群算法很方便地得到服务路径,从而确定 Web 服务路由路径。

2.1 小世界网络结构特性

^{*})十五国家科技攻关项目(编号:2002BA104C);国家 863 项目(编号:2002AA411030)。徐新卫 博士生,讲师,主要研究领域:信息系统集成和 Web 智能;丁秋林 教授,博导,研究领域:数字集成制造和系统集成。

表征小世界网络可以使用网络特征路径长度(L)和相近系数 $\gamma^{[10,11]}$ 。

定义 1 一个图的特征路径长度(L)是将图中的每个顶点 $v \in V(G)$ 连接到索引其他顶点的最短路径长度的均值的中位数。

定义 2 Γ_v 相近系数 γ_v 对于任意顶点 v , 与其邻接的顶点之间接近的程度, 即 $\gamma_v = |E(\Gamma_v)| / \binom{k_v}{2}$ 。 $|E(\Gamma_v)|$ 是 v 的领域中边的数目, $\binom{k_v}{2}$ 是 Γ_v 可能存在边的总数。 k_v 是给定子图 Γ_v 中的顶点数。对于图 G 的相近系数 γ 则为各个顶点的相近系数的平均数。

小世界网络中大多数节点的连接度都不大, 只有少数节点的连接度很高, 可以将这些少数节点看成中心节点(或称 hub 节点), 起着标识簇的作用, 同时依靠一些节点连接不同的区域(称为捷径), 它们使得网络通信范围更广, 可用资源更丰富, 查询和搜索效率更高。

2.2 构建服务网络

服务可用服务属性偏序集来对相近服务进行聚集, 由关键属性决定服务的聚集集合。这样服务就可以表示具有相近的节点簇, 选取服务簇中的一个节点作为 hub 点, 连接服务簇的所有节点, 再将 hub 点连接成为环, 并随机选取节点连接成为捷径, 就可以构建一个具有小世界属性的服务网络。在对服务进行查询、调用和执行时, 由这些服务簇构成的小世界网络具有获取服务的路径短、效率高等特点。

2.2.1 Hub 节点的选取

hub 点的选取依据是顶点的度数的多少。在每一个服务集中选取一个度数最大的顶点作为 hub 点。实际的做法是: 服务集中每个顶点随机选择顶点发送消息, 统计各个顶点收到的消息数量, 按照消息数的大小降序排序, 消息数最大的顶点为 hub 点, 排序的序列作为备选 hub 点。再将选取的 hub 点与所有顶点连接, 这样形成平等式的连接。备选 hub 点在当前 hub 节点失效时成为关键节点, 获取环路控制权和对所在的服务集进行管理。

2.2.2 连接捷径

小世界网络中, 弱连接的作用是保证在 hub 节点所构成的各个簇之间的连接, 是实现小世界网络的最小特征长度的关键, 也是 Web 服务路由路径的关键。在 Web 服务具有相近性附着性质, 即相近的服务节点可能聚集在一个 hub 节点上, 且新增的服务会依附在具有相近的 hub 节点上。弱连接则是将 hub 节点所形成的聚集连接起来, 构成具有小世界属性的网络。

按如下步骤确定连接捷径。

Step 1: 对所有服务簇产生的各个 hub 点用标号 T 标识每个 hub 点, 形成环路;

Step 2: 确定捷径数量 $s = \lceil N \cdot a \rceil$, N 为 hub 点数量, a 为系数取 $0.1 \sim 0.3$;

Step 3: 随机选取 s 个顶点, 连接到标号为 $T_j = (T_i + \lceil i/2 \rceil) \bmod N, i \in [1, s]$, 的顶点上形成捷径。 T_i 为 i 顶点的标号。

由各个对等服务在经过服务聚集、hub 点的选取和捷径的确定后, 使得网络结构具有如图 1 所示的结构。

在这种结构中, 对于每一个服务簇集, 其相近系数 $\gamma = \frac{2(N-2)}{N \cdot (N-1)}$, 当 $N > 10, \gamma < 0.2$, 其特征路径长度 $L \leq 6$, 符合

小世界网络性质。

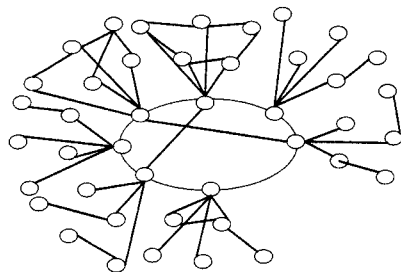


图 1 Web 服务的小世界网络模型

2.2.3 服务网络的维护

节点的连接具有偏好依附的特性, 网络规模随着新节点的加入而增大。新加入的节点偏向于连接到已存在的具有较大连接度的节点上去, 即新增的服务按照服务属性选择服务簇并连接到此服务簇的 hub 节点上。

某服务簇的 hub 节点失效时, 此 hub 点服务集中的次大连接度数的节点会获取 hub 节点前后的连接顶点信息, 升级为当前 hub 点, 接管当前服务簇中的所有有效顶点。但当前次大连接度数的节点多于服务簇的有效节点数目的二分之一时, 按 2.2.1 小节重新选择 hub 节点, 否则随机选择节点为 hub 点。

使用令牌环的管理机制, 在 hub 节点环路中判断 hub 点的状态: 在环中任意指定顶点(一定是 hub 点), 管理令牌收发令牌的收发出现故障时, 说明有 hub 点失效, 轮询定位故障 hub 点。当故障位置出现新 hub 点并有效时, 令牌继续发送。如果管理令牌的 hub 点处于失效状态, 则随机指定令牌的管理节点, 再进行令牌的轮询, 从而保证环路的健壮性和稳定性。

3 Web 路由的实现机理

WS-Routing 的显著特性是对一个逻辑名字发送消息的能力, 可以对 Web 服务实现概念化的处理方式。WS-Routing 可以实现指定消息路径的目标, 即可以通过用户的指定来确定消息传递的路径。当然, 这个路径可以是单向的, 也可以是双向的, 这样保证消息的有向性和路径可检测性。对于 WS-Routing, 可以借助 SOAP 路由的动态配置来实现。

3.1 Web 路由的逻辑名

逻辑名称是一个 URI(Unify Resource Identify), 不代表消息发送的确切位置, 可以映射到与名称所暗示的完全不同的位置, 甚至可以有多个不同的物理位置。路由感知(routing-aware)层可能是运行在与作为消息发送者相同进程内的软件, 也可能是一个作为中介服务的完全不同的服务器。在实际运用中, 在消息发送机和任何中介服务器中一般都存在路由层。图 2 示出 WSE 中路由感知层的作用。

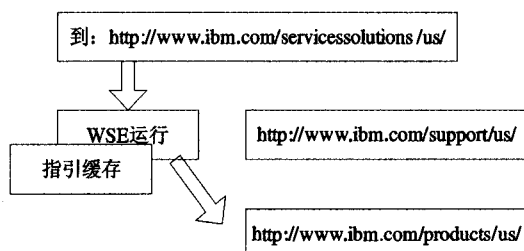


图 2 路由感知层的使用

3.2 路由消息

WS-routing 可以由用户来指定消息传递的路径,还能指定响应消息的路径。这些消息的路径放在 SOAP 的〈Header〉和〈Body〉数据容器中。默认情况下,〈Header〉发送数据时用于消息的最终目的地,当然用户可以利用 actor 属性来修改这个目的地。〈Body〉是 SOAP 消息最终目的地。WS-routing 利用表 1 的元素和属性来设置路由消息路径。

表 1 WS-routing 路由标头及元素

标头的值	含义
Path	消息路径
fwd	消息的正向路由路径
via	消息路由的中间路径
to	消息的最终目的地
from	消息的发送者
action	消息的预期目的地
rev	消息的反向路由路径

在实际创建 WS-routing 的应用程序时,通过 WSE 使用 SoapContext 的 Path 属性,在 WS-routing Path 中很方便地插入中介;在 ASP.NET 服务中使用 WSE,在虚拟目录的 Web.config 中添加如下的 HttpHandler:

```
<system.Web>
  <httpHandler>
    <add verb="*" path="*.ashx"
      type="Microsoft.Web.Services.Routing.Routing-
        Handler,
        Microsoft.Web.Services"/>
  </httpHandler>
</system.Web>
```

3.3 SOAP 路由的动态配置

WS-Routing 对于任何特定的路径缺少一种方式通知路由将消息发送到何处,因为客户端需要知道整个路由。WS-Referral 可以让路由由节点发送和接受路由配置数据,描述了一个用于路由指引的架构。这个架构可以动态配置 SOAP。

通常,WS-Referral 文档接受如下形式:

- (1) For 某个名字;
- (2) 经由某个其他名字 Go;
- (3) If 这个条件为真(如,在指引的有效期)

说明:WS-Referral 语句对 For 的描述可以是精确匹配,也可以只带前缀的,还可以包括 prefix 匹配,可用于匹配任何数量的、恰好匹配给定前缀的 URI。

3.4 基于蚁群算法的路由机制

在小世界网络中,Web 服务一旦明确了源节点和目的节点,如何确定一条路径,能够满足 QoS 要求。蚁群算法(ant colony algorithm)是模拟蚂蚁群体智能行为的优化算法,具有并行、分布式计算机制^[12]。利用蚁群算法于具有小世界网络性质的服务网络可以高效地发现一条有效路径来获取服务。

3.4.1 问题分析

可将小世界网络表示为一个有向图 $G=(V,E)$, V 是所有交换节点的集合, E 是图中所有边的集合,边表示为相邻两个节点间的直达通信路径。选取 QoS 约束指标为: B_w , D_w , L_w 和 J_w , 分别表示 QoS 要求的带宽、时延、丢失率和时延抖动限制^[13]。对于一个路由请求 w , 有下面的约束和限制:

(1) 在 w 的路由的每条路径 l 上,带宽可用限制为: $B(l) \geq B_w$;

(2) 在 w 的路由,端到端的时延限制为:

$$\sum_{n \in V_l} DN(n) + \sum_{l \in E_l} DL(l) \leq D_w;$$

(3) 在 w 的路由,端到端的丢失率限制为:

$$\prod_{n \in V_l} (1-LR(n)) \geq (1-L_w);$$

(4) 在目的节点,端到端的时延抖动限制为:

$$NDV(d) \leq J_w.$$

$B(l)$ 表示链路可用带宽, DN 表示节点处理时延, DL 表示链路时延, V_l 表示 w 路由节点集合, E_l 表示 w 路由链路的集合, LR 表示节点丢失率, NDV 表示目的节点 d 的节点时延变化。

3.4.2 基于 QoS 路由算法

在小世界网络中,根据节点所属的作用不同(一般对等节点和 hub 节点),信息素和残留信息素量也不同。

(1) 状态转移规则。位于节点 m 的第 i 只蚂蚁来选择节点 n , 依据下述规则:

当 $q < q_0$,

$$\rho_i(m, n) = \begin{cases} \max(\tau(i, m, n), & \text{若 } n \in J_i(m) \\ 0, & \text{否则} \end{cases} \quad (1)$$

否则

$$\rho_i(m, n) = \begin{cases} \frac{\tau(i, m, n)}{\sum_{u \in J_i(m)} \tau(i, m, u)}, & \text{若 } n \in J_i(m) \\ 0, & \text{否则} \end{cases} \quad (2)$$

(2) 信息素的局部调整。对于第 i 只蚂蚁,如果节点 m, n 是该蚂蚁所选路径的两个相邻节点,则信息素调整为

$$\tau(i, m, n) \leftarrow (1 - \alpha_0) \tau(i, m, n) + \alpha_0 \cdot Q \quad (3)$$

否则,信息素不变。

可以得出,如果没有局部调整,所有蚂蚁将在前一次的最好路径的有限区域内搜索。

(3) 信息素全局调整。按下述规则进行信息素的全局调整:

$$\begin{cases} \tau(i, m, n) \leftarrow (1 - \alpha_1) \tau(i, m, n) + \alpha_1 \cdot S, & \text{若 } m, n \text{ 为路径相邻节点} \\ \tau(i, m, n) \leftarrow (1 - \alpha_1) \tau(i, m, n), & \text{否则} \end{cases} \quad (4)$$

S 为限制函数,且 $S = S_2 - S_1$, 而

$$S_1 = \sum_{i=1}^V \sum_{j=1}^V LC_{ij} \cdot P_{ij}^d \quad (5)$$

若从 i 到 j 节点的边是第 d 蚂蚁所选的路径,则 $P_{ij}^d = 1$, 否则, $P_{ij}^d = 0$ 。

$$S_2 = A \cdot \sum_{i=1}^V \sum_{j=1}^V H(Z_{ij}) + B \cdot H(Z_2) + C \cdot H(Z_3) \quad (6)$$

A, B, C 表示带宽可用限制、端到端时延和端到端丢失率限制,且为正数。

如果 $Z < 0$, $H(Z) = 0$, 否则 $H(Z) = Z$ 。

$$Z_{ij} = P_{ij}^d \cdot LB_{ij} - B_w$$

$$Z_2 = D_w - (\sum \sum LD_{ij} \cdot P_{ij}^d + \sum ND_i \cdot NL_i) \quad (7)$$

$$Z_3 = \prod_{i=1}^V (1 - N_i^d \cdot NL_i) - (1 - L_w)$$

如果节点 i 是第 d 只蚂蚁所选路径的节点,则 $N_i^d = 1$, 否则为零; LB_{ij} , LC_{ij} , LD_{ij} 分别表示从节点 i 到 j 节点的边的带宽、费用和时延; ND_i , NL_i 表示节点 i 的处理时延和丢失率。

(4) 小世界网络的信息素调整。小世界网络中,如果路由的路径长度 $L_k > SD$, SD 为分离度,则 $\tau(i, n, m) = 0$, 否则不变。

基于小世界网络的 QoS 约束的路由算法步骤如下:

Step1 如果 $NDV(d) > J_w$, 则 EXIT。

Step2 删除带宽小于需求带宽的链路。

Step3 初始化网络各边和节点的信息素。

Step4 利用 m 只蚂蚁, 先从节点集合中随机选择一个节点, 按状态转换规则选择路径。如某只蚂蚁到达目的节点前死亡, 则重新释放一只蚂蚁, 重新选择从源接点到目的节点的路径。当某只蚂蚁成功完成路径的选择, 则该路径进行信息素的局部调整, 并按小世界网络的信息素调整规则来更新路径的信息素。

Step5 对所有的蚂蚁重复 step4, 直到 m 只蚂蚁都完成。

Step6 选择最小费用并满足 QoS 约束的路由, 并利用信息素的全局调整对路径进行信息素的更新。

Step7 重复 4~6 步骤, 直到获得结果。

4 算法性能分析

构造如图 3 所示的网络中, 各个顶点用 (S_y, D_s, S_b) 表示, 代表节点时延、节点丢失率和节点时延变化。边用 $(cost, Bw, L_s)$ 表示, 其中元素分别表示边的费用、带宽和链路时延。

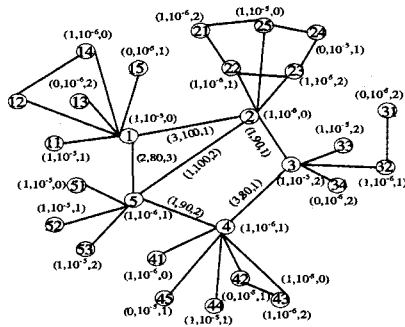


图 3 网络的拓扑结构和参数

QoS 限制要求指标和算法参数设置为: $B_w=80, D_w=12, L_w=10^{-5}, J_w=3$, 信息素的初值为 100, $m=20, \alpha_0=0.071, \alpha_1=0.078, S=0.3, q_0=0.9, A=0, B=10, C=15, SD=6$, 并对环路以外的所有边利用随机函数进行初始化, 提出路由申请, 其结果如表 2 所示。

表 2 基于蚁群的小世界网络路由结果

路由请求	基于小世界网络模型 QoS 约束下蚁群路由结果		
	选择路由	费用	时延
(43, 21)	(43→4→5→2→25→21)	7	9
(34, 52)	(34→3→2→5→52)	5	8
(14, 23)	(14→1→2→23)	5	8

表 3 仿真实验的计算结果

算法	平均迭代次数	平均运行时间(s)	平均准确率(%)
常规的基于 QoS 蚁群路由算法	327	100.3	78
基于小世界 QoS 蚁群路由算法	158	67.4	82.4

基于小世界网络模型下的 QoS 限制下的蚁群路由算法和常规蚁群路由的算法比较如表 3 所示。对于指定的路由请求, 运行 200 次, 蚂蚁数目为 30。

利用小世界网络模型的基本属性, 基于小世界的 QoS 蚁群路由算法可以提高路由效率和准确率。

5 相关工作

P2P 是基于对等计算模式, 合理运用网络资源以及最大限度实现负载均衡。由于服务的对等性, 在 P2P 网络中实现路由有一定的困难, 如泛洪通信导致链路上的广播风暴, 节点的易失性导致路由的不稳定性。对 P2P 路由的研究主要将网络结构化和非结构化方式进行。结构化的 P2P 网络基于一个覆盖网络 (Overlay network) 进行路由和查询^[14]; 非结构化的形式主要在网络拓扑上选取一个有利节点充当区域服务器, 解决系统扩展问题, 在查询和路由上采取跳步查询机制和随机漫步路由机制^[15]。

复杂网络的路由问题往往属于 NP 问题, 利用生物计算的智能性和并行处理能力来解决是当前研究的热点, 如基于蚂蚁算法和 DNA 计算。Adleman 实验就是利用 DNA 解决一个有向图的 Hamilton 路径的问题, 在连接反应中利用大量的寡核苷酸片段, 连接成边的寡核苷酸, 最终形成各种 DNA 分子, 可以看作是对应于该图的随机路径编码^[16]。

总结 根据 Web 服务的聚集依附性, 构造出具有小世界性质的服务网络模型, 使得网络具有比较小的特征路径长度和较大的相近系数, 并利用环形令牌机制对网络进行维护和管理。借助于 Web 路由的名称机制、路由消息、SOAP 动态配置和基于 QoS 约束的蚁群路由算法实现 Web 路由, 通过实验仿真, 验证在对等的小世界网络模型下 Web 服务的路由的实现机制, 得出在 QoS 约束下的蚁群路由算法具有高效的路由性能。

参考文献

- Burstein M H, Hobbs J R, Lassila O, et al. DAML-S: Web services description for the semantic Web[C]. In: Horrocks, ed. Proc of the Int'l Semantic Web Conf, Sardinia: Springer-Verlag, 2002. 348~363
- Bell D, Bussler C, Yang Jian. The Semantic Web and Web Services[J]. Information Systems, 2006, 31(4-5): 229~231
- Clark D. Next-generation Web services[J]. IEEE Internet Computing, 2002, 6(2): 12~14
- Watts D J, Strogatz S H. Collective dynamics of 'small world' networks. Nature, 1998, 392: 440~442
- Kleinberg J. The Small World phenomenon: An algorithmic perspective[D]. In: ACM Symp. on Theory of Computing, 2000
- Wang Xiaofan, Chen Guanrong. Complex networks small-world scale-free and beyond[J]. IEEE Circuits and Systems Magazine, 2003, 3(1): 6~20
- Ballinger K. The architecture of Web services[M]. 北京: 中国电力出版社, 2004. 180~191
- Sporns O. Small-world connectivity, motif composition, and complexity of fractal neuronal connections[J]. Biosystems, 2006, 85(1): 55~64
- Ferraz C H A, Herrmann H J. The Kauffman model on small-world topology[J]. Physica A: Statistical Mechanics and its Applications, In Press, Corrected Proof, Available online 17 May 2006
- Chandra A K, Dasgupta S. A small world network of prime numbers[J]. Physica A: Statistical Mechanics and its Applications, 2006, 357(3-4): 436~446
- Lahtinen J, Kertész J, Kaski K. Sandpiles on Watts-Strogatz type small-worlds[J]. Physica A: Statistical and Theoretical Physics, 2005, 349(3-4): 535~547
- 段海滨. 蚁群算法原理及其应用[M]. 北京: 科学出版社, 2005. 225~238
- Wang Xingwei, Cao Jiannong, Cheng Hui, et al. QoS multicast routing for multimedia group communications using intelligent computational methods[J]. Computer Communications, 2006, 29(12): 2217~2229
- Nejdl W, Wolpers M, Siberski W, et al. Super-peer-based routing strategies for RDF-based peer-to-peer networks[J]. Web Semantics: Science, Services and Agents on the World Wide Web, 2004, 1(2): 177~186
- Ferreira R A, Jagannathan S, Grama A. Locality in structured peer-to-peer networks[J]. Journal of Parallel and Distributed Computing, 2006, 66(2): 257~273
- Landweber L F, Kari L. The evolution of cellular computing: nature's solution to a computational problem[J]. Biosystems, 1999, 52(1-3): 3~13