(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 103338188 B
(45)授权公告日 2016.02.10

(21)申请号 201310227082.9
(22)申请日 2013.06.08
(73)专利权人 北京大学
地址 100871 北京市海淀区颐和园路5号北京大学
(72)发明人 沈晴霓 万冕 吴中海 卿斯汉
(74)专利代理机构 北京君尚知识产权代理事务所(普通合伙) 11200
代理人 余长江
(51)Int.Cl.
H04L 29/06(2006.01)
H04L 9/32(2006.01)
(56)对比文件
US 2007030973 A1,2007.02.08,
CN 101373528 A,2009.02.25,
CN 102045634 A,2044.05.04,
杨斌.基于聚类的异常检测技术的研究.《中国优秀硕士学位论文全文数据库》.2009,
审查员 段燕辉

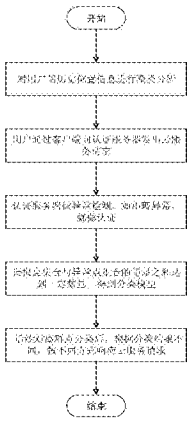
权利要求书2页 说明书9页 附图2页

(54)发明名称

一种适用于移动云的客户端动态认证方法

(57)摘要

本发明公开了一种适用于移动云的客户端动态认证方法。本方法为:1)对用户的历史位置信息进行聚类分析;2)用户通过客户端向认证服务器发出云服务请求;3)认证服务器确定该客户端的当前位置信息所对应的簇A,如果簇A中的数据点个数与总数的比值小于阈值t,则判定当前位置信息为离群点并进行认证,如果认证通过,则将其记录到一误报点集合中并响应该请求;否则拒绝该请求,并将其记录到一异常点集合中;4)如果误报点集合与异常点集合的记录之和大于阈值L,则将其标记后合并作为分类的训练数据,得到一分类模型;5)对于该用户后续的云服务请求,利用该分类模型对离群点进行分类响应云服务请求。本发明提高了用户的数据安全及账户安全。



CN 103338188 B

1. 一种适用于移动云的客户端动态认证方法，其步骤为：

1) 采用聚类分析方法对移动云用户的历史位置信息周期性进行聚类分析，分为若干簇；

2) 移动云用户通过客户端向认证服务器发出云服务请求，并且客户端将该移动云用户的账号及密码信息、地理位置信息发送给认证服务器；

3) 认证服务器收到该云服务请求之后，解析出账号、密码和位置信息，根据账号、密码对该移动云用户进行第一层验证，如果不匹配，则拒绝服务请求；否则，转到步骤 4)；

4) 认证服务器根据所述聚类分析方法确定该客户端的当前位置信息所对应的簇 A，如果簇 A 中的数据点个数与数据点总数的比值小于设定阈值 t，则判定该客户端的当前位置信息为离群点，否则为正常点并响应该请求；

5) 如果为离群点，则认证服务器对该用户进行认证，如果认证通过，则将该离群点作为一误报点并记录到一误报点集合中，并响应该请求；如果认证未通过，则拒绝该请求，并将其记录到一异常点集合中；

6) 认证服务器检查该误报点集合与异常点集合的记录之和，如果大于设定阈值 L，则将误报点和异常点进行标记后合并作为分类的训练数据；然后用分类算法对该训练数据进行分类，得到一分类模型；

7) 对于该移动云用户后续的云服务请求，认证服务器利用该分类模型对离群点进行分类，并根据分类结果确定是否响应该移动云用户的云服务请求。

2. 如权利要求 1 所述的方法，其特征在于所述认证服务器对该云服务请求进行周期性检测，判定该客户端的当前位置信息是否为离群点，直至该移动云用户退出该云服务请求。

3. 如权利要求 1 所述的方法，其特征在于认证服务器通过为误报点和异常点分别增加一“真正异常”属性，对误报点和异常点进行标记；其中，误报点该属性值为 0，异常点该属性值为 1。

4. 如权利要求 1 所述的方法，其特征在于所述聚类分析方法为 K-Means 聚类分析方法。

5. 如权利要求 1 所述的方法，其特征在于如果误报点所对应簇号簇大小与误报记录中属于该簇的点数之和大于或等于设定阈值 t，则认证服务器将所有属于该簇的误报点记录复制到正常点数据集，更新聚类训练数据集。

6. 如权利要求 1 所述的方法，其特征在于对于该移动云用户后续的云服务请求，如果认证服务器判定该客户端的当前位置信息为离群点，且该离群点与误报点集中的某数据点或某个簇的质心相似度小于设定阈值，则直接将其加入所述误报点集并响应该请求。

7. 如权利要求 1 ～ 6 任一所述的方法，其特征在于认证服务器对该移动云用户进行认证的方法为：移动云用户设置一密保手机，认证服务器将验证信息以短信的形式发送到该密保手机，移动云用户再将收到的信息发回给认证服务器。

8. 如权利要求 1 ～ 6 任一所述的方法，其特征在于认证服务器对该移动云用户进行认证的方法为：认证服务器向该移动云用户客户端发出认证请求，移动云用户将之前使用服务中的行为特征发送给认证服务器。

9. 如权利要求 1 ～ 6 任一所述的方法，其特征在于认证服务器对该移动云用户进行认证的方法为：移动云用户通过设置密保邮箱、密保令牌或密码提示问题的方式，响应认证服务器的认证请求。

10. 如权利要求 1 所述的方法，其特征在于所述异常点集合中的异常点记录信息包括经纬度、时间；所述误报点集中的误报点记录信息包括经纬度、时间及指派的簇号。

一种适用于移动云的客户端动态认证方法

技术领域

[0001] 本发明涉及一种适用于移动云的客户端动态认证方法，属于移动云环境的安全领域，主要应用在移动云计算环境下由移动客户端访问云服务的认证过程，保证用户的数据安全及账户安全。

背景技术

[0002] 移动云计算是一个将移动计算与云计算融合的产物。它希望用云计算技术存储和处理移动设备上的数据，从而缓解移动设备的固有限制，这样移动应用就能以较低的成本为用户提供更加丰富的服务体验。

[0003] 首先，应该明确一下移动云的概念。所谓移动云，移动设备与云端的计算设施共同完成数据存储和处理等计算任务的计算设施。虽然移动云与云计算有重叠的部分，但是两者还是有区别的。首先，云计算为用户提供服务时，无需让他们知道服务部署在哪里或者它们怎么被递送。移动云计算则旨在支持移动性，使用户可以用无线技术访问资源。其次，在移动设备内建立云来存储和处理数据是可能的。广义上说，目前绝大多数的移动应用都属于移动云应用，它们在设备上先完成简单的数据处理任务（文档的编辑、输入的生成等），再由云端的服务器完成复杂的计算任务（包括存储及相对复杂的计算）。而用户在使用移动云应用时，往往要提供身份信息，才访问相应的服务。然而，当前云应用的认证方案强度明显过低，并且没有考虑到设备使用者变化的情况。因此，结合移动设备的便捷性及易丢失性，本专利提出的认证方法主要针对设备丢失的情况而言的。

[0004] 目前，移动云应用在移动存储、移动商务、移动搜索、移动社交网络、即时通信等方面都有广泛的应用。除了移动搜索外，账号的安全对保证用户的数据、隐私及财产安全都是至关重要的。因此，有必要对用户做强度更高的认证。但是，由于移动设备固有的资源限制，移动云应用选用了最简单的静态口令作为用户的身份凭证。而且，类似于某些 Web 应用，移动云应用为了方便用户，基本都推荐使用“记住密码”功能。这使得原本就不安全的认证方案变得更加脆弱。此外，Web 应用的“记住密码”功能都设置了有效期，而移动云应用并没有。

[0005] 针对以上的分析，在进行云环境的集群内部数据迁移过程中，普遍存在以下安全风险：其一，由于电池等资源的限制，使得移动平台难以运行防护程序，从而无法保证移动设备的软件安全。另外，移动应用商店充斥着许多仿冒官方应用的“山寨应用”，这些应用通常都留有后门，并且极易混淆，用户极易下载到包含恶意代码的手机应用。这些因素都导致攻击者极易窃取用户的口令信息，从而伪造用户身份。其二，由于手机的便捷性，用户在使用过程中被攻击者利用窥屏等社会工程学手段获得密码的机会也更大。一旦攻击者获取了密码，就可以冒充用户访问服务。其三，目前很多应用将设备使用者作为信任度极高的实体。因此，一些实时的认证都是验证用户是否在使用指定的设备。这种实时认证能够抵御前 2 种安全风险。然而，由于移动设备的移动性更强，其丢失的风险更大，也就是说，设备随时可能丢失。一旦设备丢失，验证设备的认证方案将完全失效，用户的账号安全面临巨大的

安全威胁，用户的个人信息、存储的数据都有可能泄漏。因此，当前的移动云的认证方法都会受到以上 3 种风险的威胁。其中第 3 类的风险难度最大，危害最强。而且一旦防范了第 3 类，前 2 个也迎刃而解。本发明针对第 3 类安全风险，制定出一个动态地认证客户端的使用者身份，即设备的使用者一定是合法用户的认证方法就显得尤为重要。这个方法既要考虑用户习惯及设备资源的限制，又要能够实时动态地验证设备的使用者是否仍然是原先的合法用户。以下是目前可查到的与移动设备及认证相关的专利情况。

[0006] 申请号为 200510105150.X，发明名称为“用于网络访问的移动认证”的发明提供一种用于通过临时和 / 或一次性口令将用户认证到网络的方法。临时和 / 或一次性口令由能够通过移动电信设备被访问的服务提供商来提供。当用户调用被发送给所述服务提供商的相应的访问请求时，临时口令经请求而被提供。服务提供商检查并断言所接收的访问请求并且通过使用专用加密方法来产生临时口令。所产生的临时口令最后被传送给用户的个人移动设备。而且，移动电信设备提供 :在用户计算设备和网络之间建立基于 IP 的连接。所述移动电信设备因此提供 :建立至少两条通信链路到网络以及到用户计算设备。这样，可以通过用户的个人移动电信设备来自主地执行认证过程。在所述用户的计算设备上安装和 / 或维护涉及认证的软件因而变得多余。

[0007] 该发明虽然也是认证客户端，但是其关注的焦点是设备是否指定的，能否收到正确的口令，以此通过认证，才能访问网络。而本专利申请的焦点主要在认证使用设备的用户是否是合法的，因此认证过程不会向设备发送口令的明文形式。

[0008] 申请号为 200810027653.3，发明名称为“一种基于智能手机的移动认证系统”的发明公开了一种利用 Mobile Key 技术，结合智能手机来进行数字签名和认证，文件加密和解密的系统，它将用户密钥或证书保存在智能手机上，并且将签名 / 认证，加密 / 解密过程也在智能手机上进行。该发明主要由一台 Windows Mobile 操作系统的智能手机、桌面电脑软件模块和智能手机软件模块，首先在桌面电脑上设计和实现一个 Mobile Key 客户端，包括文件加密、解密工具，Office 签名、认证插件，再在智能手机上设计和实现一个数字签名，认证，文件加密 / 解密等数学运算的安全系统。该发明除了具有使用方便，安全性高等 USB Key 固有的优点外，还具有许多 USB Key 所不具备的优点，包括运算速度快、支持长密钥、支持处理超大文件、共享性好、安全性更高、可扩展性更好。

[0009] 该发明的主要思想是让智能手机扮演 USB Key 的作用，智能手机将完成签名 / 认证、文件加解密等操作，然后认证的信息尽管加密了，但仍然存储在手机内，因此无法抵御设备遗失的安全风险。

[0010] 申请号为 200910154847.4，发明名称为“一种基于手机 sim 卡贴片的银行业务移动认证方法”的发明公开了一种基于手机 sim 卡贴片的银行业务移动认证方法，在用户的手机 sim 卡槽中插入可存放用户证书并进行数字签名的贴片，当用户发起交易，账户资金变动时，后台交易系统向移动认证网关发送交易信息，而移动认证网关将这一交易信息发送给用户手机请求数字签名。这一请求被贴片接受后显示给用户，用户确认签名后将签名结果返回给移动认证网关，并转发给后台交易系统。后台交易系统将签名结果和请求签名内容进行验签操作，确认签名有效后对用户账户进行相应操作。应用该发明的方法，当用户资金变动时，能够及时的请求用户数字签名，保证用户资金安全，并提供用户外出时的便利性，同时具备抗抵赖性保证双方的交易安全。

[0011] 该发明没有将用户证书放在操作系统中，而是放在 sim 卡中，安全性有所提高，但仍然是与移动设备绑定。攻击者获得设备后仍然能够通过认证过程。

[0012] 申请号为 201110398800.X，发明名称为“基于音频的非接触 IC 卡及移动认证数据传输装置”的发明包括：音频接口、音频通讯编解码模块、CPU 控制模块、非接触 IC 卡模块、安全密钥处理模块和发射模块，其中：音频接口与手机相连并传输音频信号至音频通讯编解码模块，音频通讯编解码模块与 CPU 控制模块相连并传输模数转换后的数据，CPU 控制模块将数据进行逻辑判断并分别输出用于非接触式 IC 应用和安全认证的数据至非接触 IC 卡模块以及安全密钥处理模块，发射模块与非接触 IC 卡模块相连并传输 APDU 指令，安全密钥处理模块输出解密结果至 CPU 控制模块。该发明可以实现基于手机银行、手机支付、电子商务等应用的加密身份认证，其预置的非接触 IC 卡模块也可作为一张异型非接触 IC 卡，在非接终端上刷卡消费。

[0013] 该发明虽然利用用户的音频作为认证数据，能够防止不法分子的伪造。但它需要预置一个 IC 卡模块。而本专利申请对移动设备没有硬件上的特殊要求。

[0014] 申请号为 03109851.7，发明名称为“基于蓝牙技术的智能移动认证方法及其应用”的发明公开了一种智能移动认证方法，包括：智能移动基站发出一个由分组数据文件写成的读出数据的文件；智能移动单元接到这个文件后，将有效数据载荷存放在存储器的缓冲区上，读出公钥文件与存放在存储器中的密钥库进行比对。当找到一个公钥与有效载荷中的公钥一致时，则打开与该公钥对应的子存储区，读出预先写入的对应文件，并将这个文件存入缓冲区中的读操作申请文件中，与公钥文件拼装成有效载荷，经过蓝牙芯片加装识别码和数据头组成一个分组文件，通过蓝牙射频传输给蓝牙智能移动基站；蓝牙智能基站接到蓝牙智能移动单元传回的分组数据后，调用智能软件包对有效载荷进行解密处理，将解密完的数据与所存储的数据进行比对，比对一致时，发出通过认证的指令。

[0015] 该发明虽然利用蓝牙技术传输有效载荷，但是文件的加解密仍然由设备上的软件模块完成。所以，获取设备的攻击者仍然可以攻击成功。

[0016] 申请号为 200710120579.5，发明名称为“基于位置认证的电子支付系统、设备、及方法”的发明公开了一种基于位置认证的电子支付系统，包括：客户端，包括定位模块和安全加密模块，其中，定位模块用于获取用户的交易位置信息，安全加密模块用于生成加密的位置宣告信息，其中，位置宣告信息包括：交易位置信息和会话标识；定位数据库，用于存储与用户的交易记录相关的位置描述信息；服务器，用于通过将交易位置信息与定位数据库中的位置描述信息进行比较，来验证来自客户端的交易位置信息所表示的位置是否为可信位置。通过该发明，进一步降低了电子支付服务人工验证成本，同时又增加了黑客在异地发出支付定单的难度，提高了现有技术方案的安全性，提高了用户体验。

[0017] 该发明主要应用于传统主机的电子支付方法，由于主机的位置相对固定不变。因此，它在验证时只是与定位数据库中的记录做比较。不一致的话，就用密码提示问题认证用户。正因为如此，方案中的位置信息要经过加解密操作。此方案存在 2 点不足：1. 它仅对发生交易时的位置作认证；2. 仅对位置历史做记录，没有作数据分析，误报率较大，会影响用户的体验。

[0018] 申请号为 201010542659.1，发明名称为“基于移动终端地理位置异常的用户安全控制方法及装置”的发明揭示了一种基于移动终端地理位置异常的用户安全控制方法，包

括：接收移动终端发送的交易请求，所述交易请求包括该移动终端的地理位置信息；匹配所述移动终端的地理位置信息与标准地理位置信息列表；当所述移动终端的地理位置信息与标准地理位置信息列表不匹配时，提高该移动终端对应的安全控制级别。该发明还提出了相应的装置，其主要目的为提供一种基于移动终端地理位置异常的用户安全控制方法及装置，提高用户交易的安全性。

[0019] 该发明的思想是通过匹配统计分析（用户设置）得出的标准地理位置信息列表，找到位置的异常，然后就提高终端对应的安全控制。但其主要思想通过得出用户习惯的位置信息列表，然后再与之匹配，得出异常检测。此发明的也存在一些不足：1. 它仅在交易时才检测位置异常，导致它依然不能实时地认证用户的安全；2. 其标准地理位置列表是通过静态的统计分析方法获得的离散点，误报率高，无法体现用户的合法活动区域及其变更和迁移情况。

发明内容

[0020] 针对移动云应用无法抵御设备遗失所导致的攻击问题，当前的专利基本没有涉及到这一领域。但随着移动云服务的广泛应用，这一攻击导致的问题将会日益严重。因此，加强移动云的认证方法的安全强度变得迫在眉睫。本发明针对此需求，利用移动设备的位置信息作为分析内容，通过检测用户的位置异常，在维持用户的使用习惯的同时，实时地认证用户的身份，确保设备使用者的确是合法用户，从而保证用户的账号安全及数据安全。

[0021] 以下重点阐述发明中的两个要点：

[0022] 一、位置异常的检测：首先，这个方法主要是对用户使用服务时的地理位置信息（即经纬度值）在云端做数据挖掘。通过聚类分析总结出用户位置的相对集中分布的规律，再根据预先定义的异常检测算法，检测出位置异常。一旦出现位置异常，并不直接拒绝向用户提供服务，而是做附加认证，该认证信息只与用户有关，不与设备绑定。

[0023] 聚类分析技术能分析用户的位置信息，总结出不同地点之间的相似性，划分出若干个区域，这些区域就是用户地点的分布特征。根据每个区域包含数据点的个数，就可以判断出该区域该用户是否经常出入。基于聚类分析的结果，就可以判断以后用户的所处地点是否是经常活动的地方。

[0024] 地理位置信息是包含经度和纬度的二元组，而经纬度是为精确表示地球上任意位置建立的地理空间坐标系。虽然在这一坐标系统中，两点的距离并不等实际距离，但两者是成正比的，因此可以使用欧几里得距离作为相似度的衡量标准。虽然我们分析单个用户的位置数据的维度不高，但考虑到服务器必须同时为数以万计的用户服务，数据量非常巨大。而且，处理过程比较复杂的算法造成聚类的效率偏低。综合以上因素，我们选用常用的、运行效率较高的 K-Means 算法作为聚类分析算法（但不局限于这一算法，其他聚类算法同样可以），其距离定义如下：

[0025] $$dist(x,y)=\sqrt{(x_{long}-y_{long})^2+(x_{lat}-y_{lat})^2}$$

[0026] x、y 分别代表两个不同的位置，long、lat 下标分别代表经度和纬度。

[0027] 用 K-Means 算法对地理位置做聚类分析：

[0028] 1. 选择 K 个点作为初始质心。

[0029] 2. repeat

[0030] 3. 计算点与每个质心的距离，将其指派到最近的质心，形成 K 个簇。

[0031] 4. 更新每个簇的质心。

[0032] 5.until 质心不发生变化。

[0033] 在 KMeans 算法中，用误差的平方和 (Sum of the Squared Error，SSE) 作为度量聚类质量的目标函数。即每个点到所属簇的质心的距离（误差），然后计算误差的平方和。误差的平方和越小，说明聚类的质心可以更好代表簇中的点，从而聚类的效果更好。SSE 的形式化定义如下：

[0034] $$SSE=\sum_{i=1}^{K}\sum_{x\in C_i} dist(c_i,x)^2 \qquad (3\text{-}2)$$

[0035] 其中，K 指的是簇的个数，x 是指数据对象，Ci 指的是第 i 个簇，ci 指的是簇 Ci 的质心，dist 是两个对象之间的标准欧几里得距离。

[0036] 常用的离群点（异常）检测算法有 5 类：(1) 基于统计的离群点挖掘方法；(2) 基于距离的离群点挖掘方法；(3) 基于密度的离群点挖掘方法；(4) 基于聚类的离群点挖掘方法；(5) 基于偏离度的离群点挖掘方法。

[0037] 在得到聚类结果之后，做异常检测。此时移动云服务商一般只能够获得没有或部分打上了分类标签的数据集，所以只能选择无监督或半监督的离群点检测方法。移动云用户访问位置缺乏固定的规律性，不适用常用的统计分布模型，使我们无法使用基于统计的方法。虽然我们只关注全局的离群点，我们很难精确决定基于距离方法的参数，同时我们只分析地理位置这种二维数据，加上计算效率的考虑，我们排除了其他各类的方法。所以基于聚类的离群点检测方法是最符合我们要求的方法。由于离群点就是聚类算法的副产物，因此在聚类分析的基础上将算法做一定的改进，就可以用于离群点挖掘。常用的基于聚类分析的离群点挖掘方法是将远离其他聚类的小聚类看作是离群点。这一方法适用于任何聚类技术，通常对簇大小或数据点与簇中心的距离设定阈值以检测离群点。

[0038] 在使用 K-Means 聚类算法的前提下，我们定义的异常检测算法：

[0039] 异常检测算法：

[0040] 1：调用 K-Means 对用户所有正常登录的位置做聚类分析，得到 K 个簇的质心；

[0041] 2：将当前地理位置数据点按照 K-Means 算法中的规则指派给距离最近的簇，此时该簇包含的点个数不变。

[0042] 3：if 指派的簇点的个数 / 点总数＜阈值 t（默认为 1/K）；

[0043] 4：then 这个数据点是离群点（异常）

[0044] 5：else

[0045] 6：then 这个点是正常点。

[0046] 二、异常数据的处理：

[0047] 有多种原因可以造成离群点的出现，而不同的成因对提高我们的方案的准确度有重要的意义。下面是一些常见的离群点的成因：

[0048] (1) 数据测量和收集的误差。尽管民用 GPS 的定位精度可以达到 10 米左右，但是数据传输过程仍然可能造成记录值是不正确的，或者缺失的。

[0049] (2) 数据中的逻辑错误。这种情况的发生大多是管理员的误操作造成的。比如，GPS 数据中经度和纬度的值不可能超过 360。

[0050] (3) 数据的内在特性造成的异常。比如，用户出差后，访问云服务的地点，相对于出差时访问地点记录来说，就表现为一个离群点。

[0051] (4) 数据可能是陈旧的。比如，用户一年前工作调动，工作地点由北京变成了上海。那么到上海后访问服务的位置，起初仍然表现为离群点。

[0052] (5) 貌似合法的行为。例如信用卡诈骗等。

[0053] 我们知道错误的数据通过技术手段无法避免前两个原因产生的离群点。其余原因造成的离群点，则是有意义的点，但处理的方案不尽相同。其中第三、四个原因造成的离群点是用户的正常行为，不需要引起重视，要求用户验证身份；由第五个原因产生的离群点则需要引起重视，加以防范。因此，我们可以将离群点再细分为噪声点、误报点和异常点。在我们的方案中，暂时不考虑噪声点。就算要考虑噪声点，由于其与异常点类似都难以预测，可以直接划入异常点。所以，异常检测算法挖掘出的离群点只有误报点和异常点两类。误报点是指虽然被检测为异常，但其实是合法用户产生的数据点；异常点就是不是合法用户产生的数据点。

[0054] 下面我们讨论一下在方案运行的不同阶段，使用不同的处理策略，以提高或维持异常检测算法的准确率。

[0055] ①前期的处理策略

[0056] 当认证方案前期实施时，移动云提供商提取的用于聚类的训练数据，无论用户账号是否被攻击过，都认为每个数据点是正常点。得到聚类结果之后，由于 K-means 寻找质心的过程，对簇内的所有样本点在各维度求平均值，才得到质心。假如聚类的样本点有明显的离群点，就会使找到质心与实际质心位置偏差过大，使类簇发生“畸变”。这样，基于有误差的结果，异常检测算法的误差也会很大。

[0057] 因此，为了保证聚类结果的准确性，在周期性聚类之后，待判断数据点（即用户当前的位置数据，等待异常检测的判断）根据异常检测结果做不同处理。如果是正常点，直接加入到用户的数据集，当作是新的聚类训练数据集；离群点除了按误报点和异常点分开记录外，不做任何操作。这与用户因出差、旅游等原因临时到某一陌生地点并访问云服务的场景对应。

[0058] 下面我们考虑用户因为工作调动的原因，常规活动的地点发生了改变。对于这一场景的用户访问产生的误报点，如果不做任何处理，那么用户将饱受繁琐认证过程的困扰，甚至弃用服务。因此，针对这种情况，有必要某些特殊的误报点合并到正常点数据集中，以提高检测的准确性。与此同时，当误报点集中的数据点够多时，可以再次聚类。如果被检测出的离群点与误报点集中的某些数据点（或某个簇的质心）足够相似（小于某阈值），直接将它加入误报数据集，可以省去对用户的验证过程。

[0059] 综合以上 2 种情况，我们提出了前期使用的处理策略，即前期异常数据处理策略：

[0060] 1：某个点检测结果为离群点，则用密码提示问题来认证用户。如果用户回答正确，那么这个离群点是误报点；否则，就是异常点。

[0061] 2：if 这一离群点是误报点

[0062] 3：then 在误报点集中添加一条记录，包括经纬度、时间及指派的簇号（聚类结果得到的）

[0063] 4：if 该簇号簇大小 + 误报记录中属于该簇的点数≥阈值 t

[0064] 5 :then 将所有属于该簇的误报点记录复制到正常点数据集中，结束

[0065] 6 :else

[0066] 7 :then 在异常点集中添加一条异常点记录，包括经纬度、时间，结束。

[0067] ②后期的处理策略

[0068] 在方案的实施过程中，我们用密码提示问题作为认证因素。通过了这一验证环节的就是误报点 ；否则，就是异常点。随着方案的实施，异常数据就打上了分类标签，形成了可以分类的训练数据。尽管离群点的数量相对于所有位置记录比较小，但当达到一定数量后，二次挖掘这些数据，对提高认证方案的准确性，尤其在降低误报率方面具有重要意义。

[0069] 现在的问题就是将根据聚类结果检测出的离群点，分成误报点和异常点两类。因此，我们可以用现有比较成熟的分类算法做分类，总结出分类模型，从而提高异常检测的正确率，也可以改善用户体验。如果分类检测将待分类点被分类成误报点，那么就不需要做认证 ；如果被分类成异常点，那么就需要做认证。同时，这个点添加到训练数据集中。具体的后期异常数据处理策略为 ：

[0070] 1 :if 误报点与异常点的记录之和≥固定的下限（假定为 10000）

[0071] 2 :then 将误报点和异常点的记录合并后作为分类的训练数据，每条记录增加一个“真正异常”的属性。误报点此属性值为 0，异常点此属性值为 1。

[0072] 3 :then 用分类算法根据训练数据，总结出分类模型。

[0073] 4 :then 以后用分类模型判断离群点属于哪一类

[0074] 5 :else

[0075] 6 :then 执行前期异常数据处理策略。

[0076] 三、认证因素的选择 :由于我们认证的因素不能在依赖用户设备的情况下验证用户身份。所以，我们提出以下可行的认证因素 ：

[0077] ■密保手机（另一部不用作访问服务的手机）:系统将验证信息以短信的形式发送到手机，用户再将收到的信息发回给系统，从而使自己的身份得到验证。

[0078] ■密码提示问题 :用户在注册时，选择性回答一些系统生成或自己输入的问题。认证用户身份时，系统随机发送先前问题中的一个，回答正确的用户就是真正的用户。

[0079] ■密保邮箱（未在访问应用服务的手机中绑定）:系统将验证的信息以邮件的形式发送给用户。用户收到邮件后，可以正常访问服务。

[0080] ■密保令牌 :这是一个专门的用户登录的硬件，与手机客户端独立。与用于登录的硬件令牌不同，密保令牌虽然也是硬件，但不在正常登录过程中使用，只在异常情况出现时使用。

[0081] ■用户使用服务中的行为特征（除了地理位置之外）:用户在访问服务的过程中，都会做很多操作，操作及其涉及的对象都可以用于验证用户身份。比如，云储存服务就可以验证用户上次操作的文件名。为了提醒用户，可以在用户退出服务前，标识出下次登录时可能要验证的内容。

[0082] 虽然用户使用服务的行为特征，具有更高的动态性。但出于兼容性、用户友好性的考虑，我们选取密码提示问题作为认证的元素，但需要增加问题的数量，克服密码提示问题的静态性。

[0083] 与现有技术相比，本发明的积极效果为 ：

[0084] 1. 精确度得到提高。之前的相关专利都已经利用了“位置异常”的概念，但往往仅仅做统计而已。而本发明则利用数据挖掘的聚类及分类技术能够提高检测异常的精确度。

[0085] 2. 动态实时性。相关的专利只有交易发生时才检测异常，这样无法避免丢失的情况。因此，我们的方案则是周期性（例如 1 分钟为一周期）上传地理位置信息，一旦出现异常动态地验证用户身份。这样，就能更好地防范手机丢失后的风险。

[0086] 3. 可行性、兼容性好。移动云平台本来可以按需分配资源，特别适合计算量大的数据挖掘的计算任务，计算速度、效果应该比较其他专利用单一服务器的好。兼容性方面，我们认证方案充分利用了移动客户端的位置上下文，对登录等使用服务的操作没有显著影响。

附图说明

[0087] 图 1 为基于位置异常的动态认证方案流程图；

[0088] 图 2 为整个认证方案实施的模块设计图。

具体实施方式

[0089] 我们以移动云应用的登录过程来说明本发明的动态认证方案流程，但不局限于登录过程。在登录之后，我们仍将定期上传位置信息，以检测异常。一旦出现异常，执行方案流程中的验证过程。

[0090] 如图 1 所示，整个认证方案的流程如下：

[0091] ①用户通过客户端向就近的服务器发出服务请求，这一模式适用于任何一种移动云计算的连接模型。此时，由于用户尚未登录，类似于当前应用，需要认证用户。客户端将账号及密码信息（用户输入或已经和应用绑定）、连同地理位置信息打包发送给认证服务器。

[0092] ②认证服务器收到用户请求之后，解析出账号、密码和位置信息。首先做账号、密码的第一层验证。如果不匹配，拒绝服务请求，转到第⑥步。否则，转到第③步。

[0093] ③认证服务器之前已经周期性（比如，每天 1 次）地对用户正常的位置信息记录做了数据挖掘分析工作。前期运行时，那么它就读出该用户的已经周期性分析好的位置聚类结果，也就是每个簇的信息。根据聚类结果，再按照定义好的异常检测算法对当前位置做异常检测。后期运行期，那么它就将当前位置按正常位置与异常位置训练出的分类算法分类成正常或异常。如果不是异常，接受用户请求，记录这一位置，作为以后的训练数据。用户登录后，在之后的服务交互过程中客户端仍需要向服务器提供位置信息，以备周期性地异常检测使用，转到第⑤步。否则，转到第④步。

[0094] ④服务器随机产生挑战因素（密保问题）发送给用户，要求用户输入正确的响应。用户把应答内容（问题答案）发送回服务器。它再对用户的应答做判断。如果与用户设定的答案相同，则接受服务请求，记录这一位置，作为以后的训练数据。否则，拒绝服务请求，转到第⑥步。

[0095] ⑤用户在使用一段时间后会出现 2 种情况：用户是否使用完毕。如果使用完毕，转到第⑥步。如果没有使用完毕，那么仍然有 2 种情况：是否到了异常检测周期（如 1 分钟）。如果没到，用户仍可以正常使用服务，重新开始第⑤步。否则，转到第③步，做异常检测。

[0096] ⑥退出应用：包括临时切换出应用及关闭应用。

[0097] 在用户的整个使用过程都在周期性地做异常检测，导致随时都可能要额外的信息认证用户身份。随着收集到的点数量的增多，逐渐由“聚类 + 异常检测”过渡到“聚类 + 异常检测 + 分类”来判断当前位置是否是异常的。由于整个方法较短的周期认证用户的位置，出现异常就要认证用户身份，可以较好地防范手机丢失后的安全风险。

[0098] 如图 2，在具体的方案实施时，需要涉及移动客户端和认证服务器两个实体。云服务请求、响应模块与具体云服务相关，在认证方案中我们不做讨论。客户端主要负责向服务器提供认证所需的信息，根据是否需要用户输入可以分为认证交互模块、地理位置感知模块。认证服务器负责接收用户的认证信息，对信息做分析，并做出是否允许用户登录的决策。根据处理数据的方法不同，分为认证决策模块、异常检测模块、聚类分析模块和数据处理模块。

[0099] 地理位置感知模块的功能：可以通过 GPS 定位、基站定位等方法获取设备当前的地理位置信息，传递给认证交互模块。

[0100] 认证交互模块的功能：提供用户界面，接收传递过来的位置信息，负责与认证服务器的整个交互过程。

[0101] 数据维护模块负责用户的相关数据维护，包括账户密码、密码提示问题答案、聚类结果，位置信息保存等。在认证过程中，数据维护模块先从数据库中查出账号对应的正确密码，供认证决策模块做判断。如果异常情况出现，数据维护模块在用户的 4 个密码提示问题中随机抽取一个，传递给认证决策模块，用于进一步的认证。聚类模块周期性聚类后的结果，也由数据维护模块保存到服务器上。

[0102] 聚类分析模块负责对每个用户的正常位置数据集周期性做聚类分析，得到的结果由数据维护模块复杂更新。

[0103] 异常检测模块接收认证决策传递的位置信息，根据现有的聚类结果及算法，做出是否异常的判断，返回给认证决策模块。

[0104] 认证决策模块负责接收用户所有的认证信息，即账号密码及密码提示问题答案。而且，它负责决定是否要做身份的附加认证。

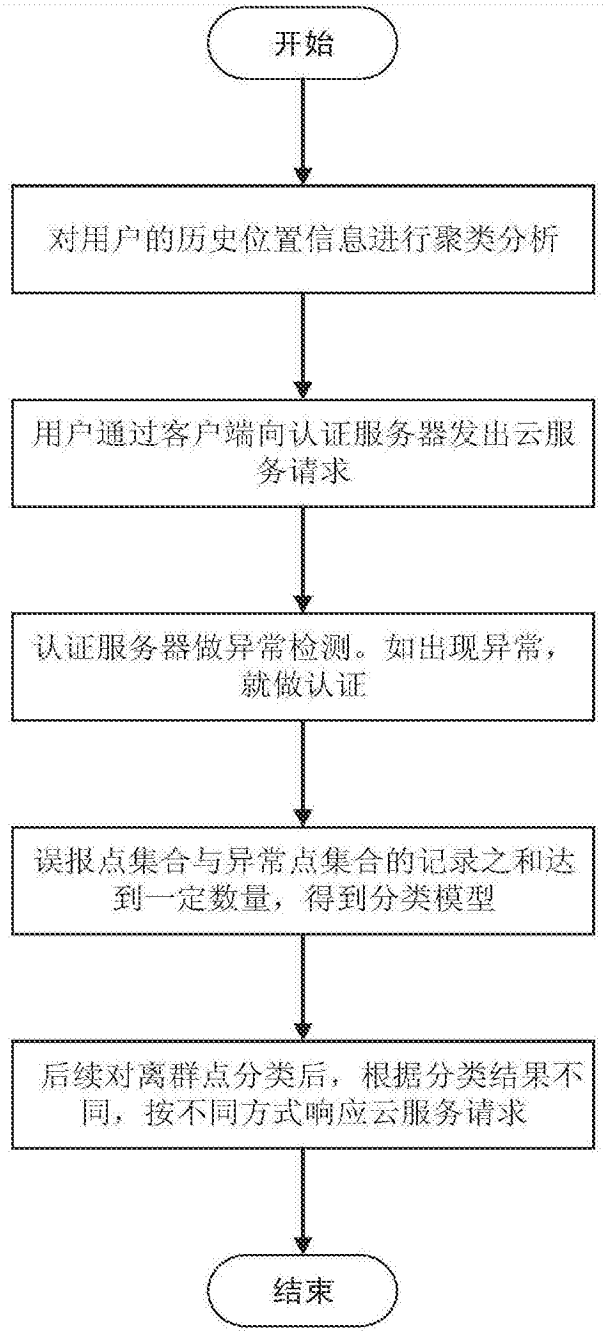

开始
对用户的历史位置信息进行聚类分析
用户通过客户端向认证服务器发出云服务请求
认证服务器做异常检测。如出现异常，就做认证
误报点集合与异常点集合的记录之和达到一定数量，得到分类模型
后续对离群点分类后，根据分类结果不同，按不同方式响应云服务请求
结束

图 1

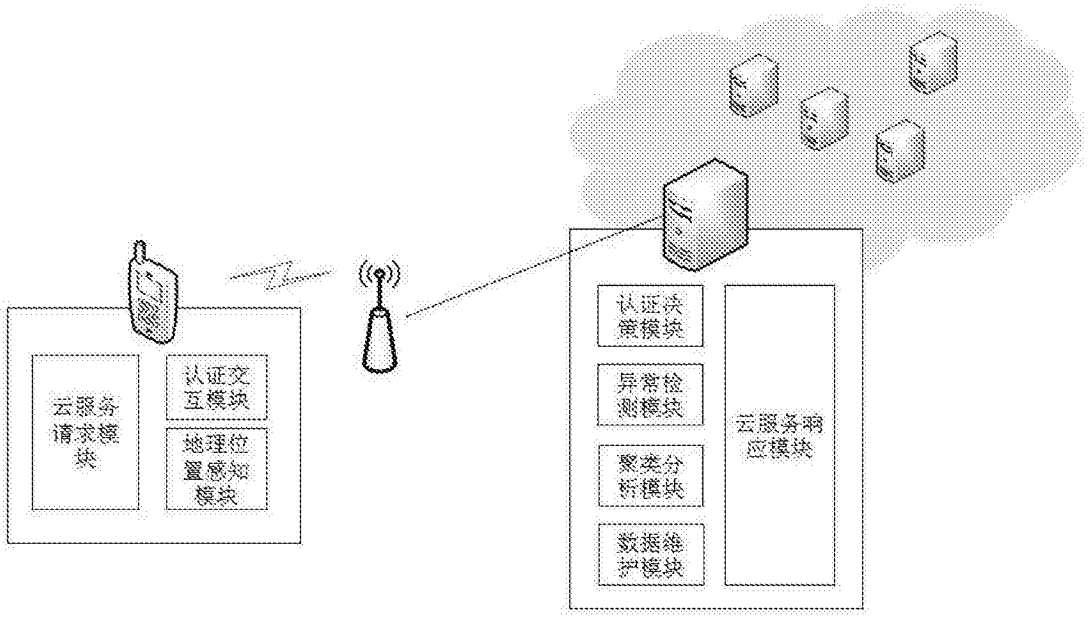

云服务请求模块
认证交互模块
地理位置感知模块
认证决策模块
异常检测模块
聚类分析模块
数据维护模块
云服务响应模块

图 2