(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 107733854 B
(45)授权公告日 2021.06.29

(21)申请号 201710761739.8

(22)申请日 2012.04.01

(65)同一申请的已公布的文献号
申请公布号 CN 107733854 A

(43)申请公布日 2018.02.23

(62)分案原申请数据
201210096275.0 2012.04.01

(73)专利权人 阿里巴巴集团控股有限公司
地址 英属开曼群岛大开曼资本大厦一座四层847号邮箱

(72)发明人 胡四海

(74)专利代理机构 北京润泽恒知识产权代理有限公司 11319
代理人 谭镇

(51)Int.Cl.
H04L 29/06(2006.01)
G06F 21/55(2013.01)
G06Q 50/00(2012.01)

(56)对比文件
CN 102200987 A,2011.09.28
CN 102339445 A,2012.02.01

审查员 许婵

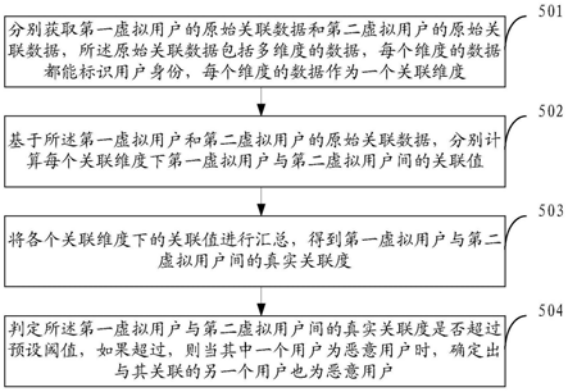
权利要求书2页 说明书12页 附图7页

(54)发明名称

一种网络虚拟账户的管理方法

(57)摘要

本申请提供了一种网络虚拟用户的风险控制方法及系统，以解决现有的方法不能准确地分析出网络虚拟用户的真实关联，进而不能准确地识别出恶意用户的问题。本申请对虚拟用户间的关联维度进行了多维度的扩展，在分析两个虚拟用户时，可以同时使用多个维度进行分析，即针对每个维度计算关联值，最后将多个维度的关联值进行汇总，得出这两个虚拟用户间最终的关联度。本申请可以准确地识别出与恶意用户关联度高的其他用户也可能为恶意用户，大大提供了网络风险控制的力度，尽可能地避免了网络欺诈行为的发生。



CN 107733854 B

1.一种网络虚拟账户的管理方法,其特征在于,包括:

分别获取第一虚拟账户的原始关联数据和第二虚拟账户的原始关联数据;

根据所述第一虚拟账户的原始关联数据和第二虚拟账户的原始关联数据,计算所述第一虚拟账户与第二虚拟账户间的关联度;

根据所述第一虚拟账户与第二虚拟账户之间的关联度,基于所述第一虚拟账户与第二虚拟账户中任意一个虚拟账户的风险状态,确定另一个虚拟账户的风险状态;

其中,所述原始关联数据包括至少一个关联维度的数据,所述关联维度的数据用于标识用户身份;

其中,在关联维度包括多个的情况下,所述关联度通过汇总多个所述关联维度下所述第一虚拟账户与所述第二虚拟账户间的关联值而获得。

2.根据权利要求1所述的方法,其特征在于,当所述原始关联数据包括多个关联维度的数据时,所述根据所述第一虚拟账户的原始关联数据和第二虚拟账户的原始关联数据,计算所述第一虚拟账户与第二虚拟账户间的关联度的步骤,包括:

针对每个关联维度的数据,计算所述关联维度下第一虚拟账户与第二虚拟账户之间的关联值;

将各关联维度下的关联值进行汇总,得到第一虚拟账户与第二虚拟账户间的关联度。

3.根据权利要求1所述的方法,其特征在于,当所述原始关联数据包括一个关联维度的数据时,所述根据所述第一虚拟账户的原始关联数据和第二虚拟账户的原始关联数据,计算所述第一虚拟账户与第二虚拟账户间的关联度的步骤,包括:

基于所述关联维度的数据,计算第一虚拟账户与第二虚拟账户之间的关联值作为关联度。

4.根据权利要求1所述的方法,其特征在于,所述根据所述第一虚拟账户与第二虚拟账户之间的关联度,基于所述第一虚拟账户与第二虚拟账户中任意一个虚拟账户的风险状态,确定另一个虚拟账户的风险状态的步骤,包括:

判定所述第一虚拟账户与第二虚拟账户间的关联度是否超过预设阈值,如果超过,则当第一虚拟账户与第二虚拟账户中任意一个虚拟账户为恶意账户时,确定出与其关联的另一个虚拟账户也为恶意账户。

5.根据权利要求1所述的方法,其特征在于,所述根据所述第一虚拟账户的原始关联数据和第二虚拟账户的原始关联数据,计算所述第一虚拟账户与第二虚拟账户间的关联度的步骤,包括:

针对所述原始关联数据中任一关联维度的数据,基于所述关联维度下的数量因素、时间因素、级联因素其中至少一项,计算第一虚拟账户和第二虚拟账户之间的关联值;

根据所述关联值计算所述第一虚拟账户与第二虚拟账户间的关联度。

6.根据权利要求5所述的方法,其特征在于,所述数量因素采用求和函数。

7.根据权利要求6所述的方法,其特征在于:所述时间因素和级联因素都采用倒数函数。

8.根据权利要求7所述的方法,其特征在于,通过以下公式计算每个关联维度下第一虚拟账户和第二虚拟账户之间的关联值:

$\Sigma_x \Sigma_{level} \Sigma_t (1/t)*(1/level)$;

其中，Σ表示求和函数，x表示某个关联维度下的关联数量，level表示级联层次，t表示时间。

9.根据权利要求1所述的方法，其特征在于，所述原始关联数据包括：

IP、cookie、机器指纹、手机号码、电话号码、传真、电子邮箱，和/或地址、登录账号其中至少一个维度的数据。

10.根据权利要求1所述的方法，其特征在于，所述分别获取第一虚拟账户的原始关联数据和第二虚拟账户的原始关联数据的步骤，包括：

获取第一虚拟账户的原始关联数据；

基于所述第一虚拟账户的原始关联数据，查找与第一虚拟账户关联的第二虚拟账户，以及所述第二虚拟账户的原始关联数据。

11.根据权利要求10所述的方法，其特征在于，所述查找与第一虚拟账户关联的第二虚拟账户，包括：

对原始关联数据中至少一个关联维度的数据，根据第一虚拟账户的标识查找该账户使用的维度数据；

利用所述查找到的维度数据，继续查找其中每个维度数据对应的账户列表；

将每个关联维度下查找到的账户列表进行去重整理，最后得到的账户列表中标识的所有账户作为与所述第一虚拟账户关联的第二虚拟账户。

12.根据权利要求2所述的方法，其特征在于，所述将各关联维度下的关联值进行汇总，包括：

将各个关联维度下的关联值进行求和。

13.一种网络虚拟账户的管理方法，其特征在于，包括：

分别获取第一虚拟账户的原始关联数据和第二虚拟账户的原始关联数据，所述原始关联数据包括至少一个关联维度的数据；

计算至少一个关联维度下第一虚拟账户与第二虚拟账户间的关联度，包括：基于数量因素、时间因素、级联因素其中至少一个，计算第一虚拟账户和第二虚拟账户间的关联度；

根据所述第一虚拟账户与第二虚拟账户之间的关联度，基于所述第一虚拟账户与第二虚拟账户中任意一个虚拟账户的风险状态，确定另一个虚拟账户的风险状态；其中，所述关联维度的数据用于标识用户身份；

其中，在关联维度包括多个的情况下，所述关联度通过汇总多个所述关联维度下所述第一虚拟账户与所述第二虚拟账户间的关联值而获得。

一种网络虚拟账户的管理方法

技术领域

[0001] 本申请涉及网络技术，特别是涉及一种网络虚拟用户的风险控制方法及系统。

背景技术

[0002] 随着互联网的发展，越来越多的人通过网络进行沟通和交流，网络已经成为众多用户的一个信息交流平台。在网络中，每一个用户都是一个虚拟用户。虚拟用户在网络中的行为在一定程度上可以反映出真实世界中用户之间的关系。

[0003] 例如，以社交网络(SNS，Social Networking Services)为例，如图1所示，如果虚拟用户A和虚拟用户B在网络中都有一个共同的好友用户C，那么用户A和用户B在真实世界中就很有可能也是好友。

[0004] 实际应用中，利用虚拟用户之间的这种真实关联，可以对用户的网络行为进行风险控制，避免网络欺诈行为。例如，在网上交易系统中，如果分析出某个用户在交易过程中存在欺诈行为，并且已经将该用户列入恶意用户的黑名单，如果能够分析出与该恶意用户有真实关联的其他虚拟用户，这些其他虚拟用户存在欺诈的可能性也很大，那么可以将这些有关联的其他虚拟用户也提前设定为恶意用户，从而尽早地避免交易欺诈的发生。

[0005] 在上述风险控制的过程中，现有技术中，一般通过机器指纹来分析网络虚拟用户的真实关联。通过采集机器数据(即机器指纹)，如硬盘、主板等能够唯一标识一台机器的数据，可判断虚拟用户是否使用同一台物理机器。如果两个虚拟用户共同使用一台物理机器，那么这两个虚拟用户可能存在关联。

[0006] 上述现有技术存在的缺点是：只有在不同的虚似用户使用同一台物理机器的时候，才能判断出虚似用户为关联用户，如果虚拟用户使用了不同的物理机器，即使他们为关联用户，上述方法也无法分析出来。因此，现有的这种分析方法太过局限，并不能准确地分析出网络虚拟用户的真实关联，进而不能准确地识别出恶意用户，为网络欺诈行为的发生提供了可乘之机。

发明内容

[0007] 本申请提供了一种网络虚拟用户的风险控制方法及系统，以解决现有的方法不能准确地分析出网络虚拟用户的真实关联，进而不能准确地识别出恶意用户的问题。

[0008] 为了解决上述问题，本申请公开了一种网络虚拟账户的管理方法，包括：

[0009] 分别获取第一虚拟账户的原始关联数据和第二虚拟账户的原始关联数据；

[0010] 根据所述第一虚拟账户的原始关联数据和第二虚拟账户的原始关联数据，计算所述第一虚拟账户与第二虚拟账户间的关联度；

[0011] 根据所述第一虚拟账户与第二虚拟账户之间的关联度，基于所述第一虚拟账户与第二虚拟账户中任意一个虚拟账户的风险状态，确定另一个虚拟账户的风险状态。

[0012] 本申请还公开了一种网络虚拟账户的管理方法，包括：

[0013] 分别获取第一虚拟账户的原始关联数据和第二虚拟账户的原始关联数据，所述原

始关联数据包括至少一个关联维度的数据;计算至少一个关联维度下第一虚拟账户与第二虚拟账户间的关联度,包括:基于数量因素、时间因素、级联因素其中至少一个,计算第一虚拟账户和第二虚拟账户间的关联度;根据所述第一虚拟账户与第二虚拟账户之间的关联度,基于所述第一虚拟账户与第二虚拟账户中任意一个虚拟账户的风险状态,确定另一个虚拟账户的风险状态。

[0014] 与现有技术相比,本申请包括以下优点:

[0015] 首先,本申请对虚拟用户间的关联维度进行了多维度的扩展,IP、cookie、机器指纹、手机号码、电话号码、传真、电子邮箱、地址、登录账号,等等能够标识用户身份的信息,都可以作为一个关联维度来分析用户间的真实关联情况。基于此,本申请在分析两个虚拟用户时,可以同时使用至少一个维度进行分析,得出这两个虚拟用户间最终的关联度。这种关联维度的扩展,在虚拟用户不使用同一台物理机器的时候,也能利用其它维度的信息分析出用户间的真实关联,因此打破了传统分析方法的限制,能够更加准确地分析出网络虚拟用户的真实关联。进而,通过这种准确的分析方法,可以准确地识别出与恶意用户关联度高的其他用户也可能为恶意用户,大大提供了网络风险控制的力度,尽可能地避免了网络欺诈行为的发生。

[0016] 其次,本申请还对关联的同一维度的分析因素进行了扩展。在对每一个维度进行关联分析时,不仅仅考虑一个因素,而是同时结合了多个因素(数量因素、时间因素和级联因素)综合分析,并汇总了同维度下各个因素对关联的影响。这种同一维度下多因素的扩展,相比现有技术中的单一因素分析,能够使分析结果更加准确。

[0017] 当然,实施本申请的任一产品不一定需要同时达到以上所述的所有优点。

附图说明

[0018] 图1是现有技术中社交网络中SNS好友的示意图;

[0019] 图2是本申请实施例中数量因素的示意图;

[0020] 图3.1至3.3是本申请实施例中时间因素的示意图;

[0021] 图4是本申请实施例中级联因素的示意图;

[0022] 图5是本申请实施例所述一种网络虚拟用户的风险控制方法的流程图;

[0023] 图6是本申请另一优选实施例中获取第一虚拟用户和第二虚拟用户的原始关联数据的流程图;

[0024] 图7是本申请实施例中用户间通过多维度关联的示意图;

[0025] 图8是本申请实施例中关联度计算的示意图;

[0026] 图9是本申请实施例所述一种网络虚拟账户的风险控制方法的流程图;

[0027] 图10是本申请实施例所述另一种网络虚拟账户的风险控制方法的流程图;

[0028] 图11是本申请实施例所述一种网络虚拟用户的风险控制系统结构图

具体实施方式

[0029] 为使本申请的上述目的、特征和优点能够更加明显易懂,下面结合附图和具体实施方式对本申请作进一步详细的说明。

[0030] 本申请提出的网络虚拟用户的风险控制方法,在分析用户真实关联时,对关联的

维度和同一维度下的分析因素都进行了扩展。

[0031] 其中，关联的维度不局限于机器指纹，IP、cookie、手机号码、电话号码、传真、电子邮箱、地址、登录账号，等等能够标识用户身份的信息，都可以作为一个关联维度来分析用户间的真实关联情况。

[0032] 而且，同一维度下的分析因素扩展出数量因素、时间因素和级联因素。

[0033] 参照图2所示，是本申请实施例中数量因素的示意图。

[0034] 数量因素是指：如果虚拟用户A和虚拟用户B都有共同的一批好友C，D,E…，那么A和B为好友的可能将会大大增加。

[0035] 参照图3.1至3.3所示，是本申请实施例中时间因素的示意图。

[0036] 时间因素是指：如果虚拟用户A和虚拟用户B都有共同的一个好友C，且都在最近一年内成为好友，那A和B是好友的可能极大，如图3.1所示；同样的，若虚拟用户A和虚拟用户C在十年前是好友，虚拟用户B和虚拟用户C在十年前也是好友，那么A和B是好友的可能也极大，如图3.2所示；反之，若虚拟用户A和虚拟用户C在十年前是好友，而虚拟用户B和虚拟用户C最近一年才是好友，那么A和B为好友的可能性就会小一些。

[0037] 参照图4所示，是本申请实施例中级联因素的示意图。

[0038] 级联因素是指：朋友的朋友的朋友…也可能是朋友，当然级联导数越多，可能性就会降低。例如虚拟用户A和虚拟用户B是好友，虚拟用户B和虚拟用户C是好友，那么A和C是好友的可能极大；进一步地，虚拟用户C 和虚拟用户D是好友，那么A和D是好友的可能下降；再进一步，虚拟用户D和虚拟用户E是好友，那么A和E是好友的可能再下降。其中，称A 和B一层关联，A和C二层关联，A和D三层关联，A和E四层关联。

[0039] 基于以上内容，下面通过实施例对本申请所述的方法进行详细说明。

[0040] 参照图5所示，是本申请实施例所述一种网络虚拟用户的风险控制方法的流程图。

[0041] 下面以分析两个虚拟用户之间真实关联的过程为例，步骤如下：

[0042] 步骤501，分别获取第一虚拟用户的原始关联数据和第二虚拟用户的原始关联数据，所述原始关联数据包括多维度的数据，每个维度的数据都能标识用户身份，每个维度的数据作为一个关联维度；

[0043] 如前所述，所述多维度的原始关联数据可以包括：IP，cookie，机器指纹，手机号码，电话号码，传真，电子邮箱，地址，登录账号，等等信息。其中每个维度的数据都能够标识虚拟用户的身份，因此每个维度的数据都可以用来分析关联度。此外，其他能够标识虚拟用户身份的信息也可以作为原始关联数据使用，在此不一一列举。

[0044] 步骤502，基于所述第一虚拟用户和第二虚拟用户的原始关联数据，分别计算每个关联维度下第一虚拟用户与第二虚拟用户间的关联值；

[0045] 例如，第一虚拟用户和第二虚拟用户的原始关联数据都包含IP和 cookie，先利用IP计算这两个虚拟用户间的关联值1，然后再利用cookie计算这两个虚拟用户间的关联值2。

[0046] 优选地，如前所述，同一维度下的分析因素扩展出数量因素、时间因素和级联因素。因此，在每个关联维度下，都可基于数量因素、时间因素和级联因素，计算第一虚拟用户和第二虚拟用户间的关联值。

[0047] 具体的基于每个分析因素计算关联值的方法，将在下面的实施例中详细介绍。

[0048] 步骤503,将各个关联维度下的关联值进行汇总,得到第一虚拟用户与第二虚拟用户间的真实关联度;

[0049] 所述汇总的方法有很多,可以根据实际应用情况而定,下面仅列举一种汇总方法,但本申请的保护范围不应限定于此。

[0050] 例如,可以将各个关联维度下的关联值进行求和,求和得到的结果即表示第一虚拟用户与第二虚拟用户间的真实关联度。其中,所述求和可以是简单的求和函数,也可以是平方求和或加权求和等函数。

[0051] 例如:

[0052] 求和=x1+x2+x3;

[0053] 平方求和=x1^2+x2^2+x3^2;

[0054] 加权求和=a*x1+b*x2+c*x3;

[0055] 注x1^2=x1*x1,a、b、c均为加权系数。

[0056] 步骤504,判定所述第一虚拟用户与第二虚拟用户间的真实关联度是否超过预设阈值,如果超过,则当其中一个用户为恶意用户时,确定出与其关联的另一个用户也为恶意用户。

[0057] 在网络风险控制中,经过步骤501至503的处理,可以得到两个虚拟用户间的真实关联度,基于此,如果这两个用户的真实关联度超过预设的阈值,表明这两个用户在真实世界中的关联度很高,因此,如果其中一个用户已经确定为恶意用户,那么与其关联的关联度高的用户是恶意用户的可能性也非常大,所以将关联度超过预设阈值的关联用户也确定为恶意用户。

[0058] 综上所述,在分析两个虚拟用户时,可以同时使用多个关联维度的数据 (如IP、cookie、机器指纹等等)进行分析。这种关联维度的扩展,在虚拟用户不使用同一台物理机器的时候,也能利用其它维度的信息分析出用户间的真实关联,因此打破了传统分析方法的限制,能够更加准确地分析出网络虚拟用户的真实关联。进而,通过这种准确的分析方法,可以准确地识别出与恶意用户关联度高的其他用户也可能为恶意用户,大大提供了网络风险控制的力度,尽可能地避免了网络欺诈行为的发生。

[0059] 基于图5实施例,在本申请的另一优选实施例中,上述步骤501可通过图6所示的步骤实现,具体如下。

[0060] 参照图6所示,是本申请另一优选实施例中获取第一虚拟用户和第二虚拟用户的原始关联数据的流程图。

[0061] 步骤601,获取第一虚拟用户的原始关联数据;

[0062] 假设第一虚拟用户为用户A,用户A的原始关联数据包括IP和cookie。

[0063] 步骤602,基于所述第一虚拟用户的原始关联数据,查找与第一虚拟用户关联的第二虚拟用户,以及所述第二虚拟用户的原始关联数据。

[0064] 以IP关联维度为例,为了找出虚拟用户A和其他所有用户共同使用过的IP的情况,去除时间因素,数据存储格式如下表1和表2:

[0065]

用户	IP列表
A	IP1,IP2,IP3
B	IP2

C	IP4

[0066] 表1

[0067]

IP值	用户列表
IP1	A
IP2	B,C
IP3	A,C
IP4	C

[0068] 表2

[0069] 上述这种key-value的存储格式可以快速查询出所有与虚拟用户A共同使用过IP的其他虚拟用户。因此,找出和A的关联用户就很简单了,步骤 602包含的子步骤如下:

[0070] 子步骤6021,对每个关联维度,根据第一虚拟用户的标识查找该用户使用的维度数据;

[0071] 仍以IP关联维度为例,查询上述表1,可以得到表3的查询结果,如下:

[0072]

用户	IP列表
A	IP1,IP2,IP3

[0073] 表3

[0074] 由表3可知,用户A使用过的IP为IP1、IP2和IP3。

[0075] 子步骤6022,利用所述查找到的维度数据,继续查找其中每个维度数据对应的用户列表;

[0076] 仍以IP关联维度为例,根据表3中的IP1、IP2和IP3,查询表2,可以得到表4的查询结果,如下:

[0077]

IP值	用户列表
IP1	A
IP2	B,C
IP3	A,C

[0078] 表4

[0079] 子步骤6023,将每个关联维度下查找到的用户列表进行去重整理,最后得到的用户列表中标识的所有用户即为与所述第一虚拟用户关联的第二虚拟用户。

[0080] 如表4所示,使用IP1的用户为A,使用IP2的用户为B和C,使用IP3 的用户为A和C。经过去重整理,最后得到的用户列表中包含用户A、B和 C。

[0081] 由此可知,一层关联就是A、B和C。去掉A自身后,就是B和C,即与用户A一层关联的用户为B和C。

[0082] 继续查找与用户A二层关联的用户,查询表1,得到表5的查询结果,如下:

[0083]

B	IP2
C	IP4

[0084] 表5

[0085] 查询表2,得到表6的查询结果,如下:

[0086]

IP2	B,C

IP4	C

[0087] 表6

[0088] 去重后，二层关联仍是B和C。

[0089] 类似的，其他原始关联数据，如cookie、机器指纹、手机号码、…，都可以采用上述的数据存储格式。

[0090] 综上所述，由图6可以看出，为了节省计算量，可以先确定一个虚拟用户，然后再查询与该用户关联的其他用户，再进行真实关联度的计算。当然，根据实际应用的需要，也可以将任意两个用户进行组合，计算他们的关联度，但可能存在关联度为0的组合情况。

[0091] 基于以上内容，在查找到用户A及其关联用户B和C之后，下面通过另一实施例详细说明如何计算用户A与其关联用户的真实关联度。

[0092] 仍以IP关联维度为例，如果用户A和B都使用过相同的IP:IP1，那么 A和B就很有可能“有关联”。在这种关联上考虑数量因素、时间因素、级联因素，如下：

[0093] 数量因素：如果A和B都用过相同的一批IP:IP1，IP2，IP3…，那么A 和B“有关联”的可能将会大大增加。

[0094] 时间因素：如果A和B都使用过相同的IP:IP1，且都在最近一年内使用过，那A和B“有关联”的可能极大；反之，若A在十年前使用过IP1，而B在最近一年才使用过IP1，那么A和B“有关联”的可能性就会小一些。

[0095] 级联因素：如果A和B使用过相同的IP:IP1，B和C使用过相同的IP:IP2。那么A和C也有可能“有关联”。可以定义：称A和B通过IP一层关联； A和C通过IP二层关联。

[0096] 综合以上三个因素，考虑“时间”和“级联层次”对“关联”的衰减效应，可以采用衰减函数进行计算；考虑“数量”对“关联”的累积效应，可以采用累积函数进行计算。对于具体的衰减函数和累积函数，本申请不进行限定。

[0097] 一种函数方案如下：对“数量”的累积函数使用求和函数，对“时间”和“级联层次”的衰减函数使用倒数函数。

[0098] 仍然以IP为例，说明如下：

[0099] 将时间(t)按月为单位分类：

[0100] t＝1代表本月，t＝2代表上个月…定义关联度函数为1/t。

[0101] 级联层次(level)：

[0102] level＝1代表通过IP一层关联，level＝2代表通过IP二层关联…定义关联度函数为1/level。

[0103] 举例如下：若时间间隔为t，级联层次为level，则关联度为(1/t)*(1/level)。

[0104] 如果A和B在本月(t＝1)，有通过IP一层关联(level＝1)(使用过相同IP1)，那么关联度为(1/1)*(1/1)＝1；

[0105] 如果A和B在上个月(t＝2)，有通过IP一层关联(level＝1)，那么关联度为(1/2)*(1/1)＝0.5；

[0106] 如果A和B在上个月(t＝2)，有通过IP二层关联(level＝2)，那么关联度为(1/2)*(1/2)＝0.25。

[0107] 综上所述，假设限制时间为半年，级联层数为3层，那么针对某个IP1，某二人A和B的关联度为：

[0108] $\Sigma_{level(level=1\sim3)}\Sigma_{t(t=1\sim6)}(1/t)*(1/level)$。

[0109] 对数量进行汇总:对数量的累积函数使用求和函数。即求和每个IP的关联度。假设我们限制时间为半年,级联层数为3层,那么针对所有IP,某二人A和B的关联度为:

[0110] $\Sigma_{ip}\Sigma_{level(level=1\sim3)}\Sigma_{t(t=1\sim6)}(1/t)*(1/level)$。

[0111] 需要说明的是,以上是以IP为例,即只考虑了一个维度(IP)。实际上,可以有更多的维度:IP,cookie,机器指纹,手机号码,电话号码,传真,电子邮箱,地址,登录账号,…,如图7所示。

[0112] 对每个维度都进行类似于IP关联的计算,如图8所示,每个维度的计算都考虑时间因素、数量因素和级联因素。最后将每个维度的计算结果汇总 (如求和)后,即可以得到网络虚拟用户之间的真实的关联性。

[0113] 例如,用户A和B之间通过IP、cookie、机器指纹、电话号码关联,计算时每个维度都考虑时间因素、数量因素和级联因素,分别计算出A和B 之间的IP关联值、cookie关联值、机器指纹关联值、电话号码关联值,然后将这些关联值求和汇总,得到A和B最终的真实关联度。

[0114] 再例如,用户A和B之间通过IP关联,B和C之间通过cookie关联,最后汇总时就可以将A和B之间的IP关联值与B和C之间的cookie关联值相加求和,得到A和C之间的真实关联度。

[0115] 此外,还需要说明的是,对“数量”的累积函数包括但不限定于上述列举的求和函数,还可以是加权求和函数或平方求和函数,或者是其他的求和方式,这些都可统称为求和函数。同样,对“时间”和“级联层次”的衰减函数包括但不限定于上述列举的倒数函数,还可以是加权倒数函数或平方倒数函数,或者是其他的求和方式,这些都可统称为倒数函数。

[0116] 其中,加权求和与加权倒数的举例如下:

[0117] 比如求和=x1+x2+x3;

[0118] 加权求和=a*x1+b*x2+c*x3;

[0119] 倒数=1/x1+1/x2+1/x3;

[0120] 加权倒数=a/x1+b/x2+c/x3;

[0121] 注:a、b、c均为加权系数。

[0122] 平方求和与平方倒数的举例如下:

[0123] 比如求和=x1+x2+x3;

[0124] 平方求和=x1^2+x2^2+x3^2;

[0125] 倒数=1/x1+1/x2+1/x3;

[0126] 平方倒数=1/(x1^2)+1/(x2^2)+1/(x3^2);

[0127] 注:x1^2=x1*x1。

[0128] 综上所述,这种对关联维度和同一维度下分析因素的扩展,在虚拟用户不使用同一台物理机器的时候,也能利用其它维度的信息分析出用户间的真实关联,因此打破了传统分析方法的限制,能够更加准确地分析出网络虚拟用户的真实关联。而且,相比现有技术中的单一因素分析,进一步使分析结果更加准确。

[0129] 上述这种分析网络虚拟用户真实关联的方法,可以应用到多种场景中。例如,在网络风险控制的场景中,如果获知某个用户具有欺诈行为不可靠,那么与该用户有真实关联

的其他用户也可能存在这种欺诈的风险。再例如，在智能推荐的场景中，如果获知某个用户喜好购买户外用品，那么也可以向与该用户真实关联的其他用户自动推荐户外用品，以提高购买度。

[0130] 参照图9所示，是本申请另一实施例中一种网络虚拟账户的风险控制方法的流程图。

[0131] 步骤710，分别获取第一虚拟账户的原始关联数据和第二虚拟账户的原始关联数据；

[0132] 在本发明实施例中，虚拟账户也可以理解为用户在网络中使用的账户，比如淘宝账户等，也可以理解为前述实施例中的虚拟用户。

[0133] 本申请实施例可以获取虚拟账户的至少一个维度的原始关联数据。其中至少一个维度的原始关联数据包括IP，cookie，机器指纹，手机号码，电话号码，传真，电子邮箱，地址，登录账号等等其中一个或多个。

[0134] 优选的，步骤710包括：

[0135] 子步骤M11，获取第一虚拟账户的原始关联数据；

[0136] 子步骤M12，基于所述第一虚拟账户的原始关联数据，查找与第一虚拟账户关联的第二虚拟账户，以及所述第二虚拟账户的原始关联数据。

[0137] 子步骤M11和子步骤M12参照前述步骤601和602，其原理类似在此不再详述。

[0138] 优选的，所述查找与第一虚拟账户关联的第二虚拟账户，包括：

[0139] 子步骤M121，对原始关联数据中至少一个关联维度的数据，根据第一虚拟账户的标识查找该账户使用的维度数据；

[0140] 子步骤M122，利用所述查找到的维度数据，继续查找其中每个维度数据对应的账户列表；

[0141] 子步骤M123，将每个关联维度下查找到的账户列表进行去重整理，最后得到的账户列表中标识的所有账户作为与所述第一虚拟账户关联的第二虚拟账户。

[0142] 子步骤M121至子步骤M123参照前述子步骤6021至子步骤6023，其原理类似，在此不再详述。

[0143] 步骤720，根据所述第一虚拟账户的原始关联数据和第二虚拟账户的原始关联数据，计算所述第一虚拟账户与第二虚拟账户间的关联度；

[0144] 在本申请实施例中，获取到第一虚拟账户的原始关联数据和第二虚拟账户的原始关联数据后，即可根据两者的原始关联数据计算两个虚拟账户之间的关联度。

[0145] 优选的，当所述原始关联数据包括多个关联维度的数据时，步骤720，包括：

[0146] 步骤M21，针对每个关联维度的数据，计算所述关联维度下第一虚拟账户与第二虚拟账户之间的关联值；

[0147] 在本申请实施例中，在有多个关联维的数据时，对于每个关联维度，都计算该关联维度下两个虚拟账户之间的关联值。

[0148] 步骤M22，将各关联维度下的关联值进行汇总，得到第一虚拟账户与第二虚拟账户间的关联度。

[0149] 然后将各个维度下的关联值进行汇总，得到两个虚拟账户之间的最终的关联度。

[0150] 优选的，所述将各关联维度下的关联值进行汇总，包括：将各个关联维度下的关联

值进行求和。当然可以用前述多种方式求和。

[0151] 优选的，当所述原始关联数据包括一个关联维度的数据时，步骤720，包括：

[0152] 步骤M23，基于所述关联维度的数据，计算第一虚拟账户与第二虚拟账户之间的关联值作为关联度。

[0153] 当只有一个关联维度数据时，可以直接计算两个虚拟账户之间的关联值，然后将该关联值作为两个虚拟账户之间的关联度。

[0154] 需要说明的是，在上述计算关联值的过程中，可以于所述关联维度下的数量因素、时间因素、级联因素其中至少一项，计算第一虚拟账户和第二虚拟账户之间的关联值。

[0155] 优选的，步骤720包括：

[0156] 子步骤M24，针对所述原始关联数据中任一关联维度的数据，基于所述关联维度下的数量因素、时间因素、级联因素其中至少一项，计算第一虚拟账户和第二虚拟账户之间的关联值；

[0157] 如前所述，同一维度下的分析因素扩展出数量因素、时间因素和级联因素。因此，在任一关联维度下，都可基于数量因素、时间因素和级联因素，计算第一虚拟用户和第二虚拟用户间的关联值。

[0158] 子步骤M25，根据所述关联值计算所述第一虚拟账户与第二虚拟账户间的关联度。

[0159] 当有一个关联维度下的关联值，可以直接将该关联值作为两个虚拟账户之间的关联度。当有多个关联维度下的关联值，可以将该关联值进行累加作为两个虚拟账户之间的关联度。

[0160] 优选的，所述数量因素采用求和函数。

[0161] 优选的，所述时间因素和级联因素都采用倒数函数。

[0162] 优选的，通过以下公式计算每个关联维度下第一虚拟账户和第二虚拟账户之间的关联值：

[0163] $\Sigma_x \Sigma_{level} \Sigma_t (1/t)*(1/level)$；

[0164] 其中，Σ表示求和函数，x表示某个关联维度下的关联数量，level表示级联层次，t表示时间。

[0165] 优选的，所述原始关联数据包括：

[0166] IP、cookie、机器指纹、手机号码、电话号码、传真、电子邮箱，和/或地址、登录账号其中至少一个维度的数据。

[0167] 上述求关联值的过程参考前述实施例的关联值的计算过程，在此不再详述。

[0168] 步骤730，根据所述第一虚拟账户与第二虚拟账户之间的关联度，基于所述第一虚拟账户与第二虚拟账户中任意一个虚拟账户的风险状态，确定另一个虚拟账户的风险状态。

[0169] 由于通过前述步骤，两个虚拟账户之间具有了关联度，那么即可在该关联度达到设定条件时，对两个虚拟账户之间的风险状态进行同步，比如其中一个有风险，那么另外一个也有风险。

[0170] 优选的，所述步骤730，包括：

[0171] 子步骤M31，判定所述第一虚拟账户与第二虚拟账户间的关联度是否超过预设阈值，如果超过，则当第一虚拟账户与第二虚拟账户中任意一个虚拟账户为恶意账户时，确定

出与其关联的另一个虚拟账户也为恶意账户。

[0172] 本步骤参照前述步骤504的描述，其原理类似，在此不再详述。

[0173] 综上所述，在分析两个虚拟用户时，可以使用一个或多个关联维度的数据(如IP、cookie、机器指纹等等)进行分析。打破了传统分析方法的限制，能够更加准确地分析出网络虚拟用户的真实关联。进而，通过这种准确的分析方法，可以准确地识别出与恶意用户关联度高的其他用户也可能为恶意用户，大大提供了网络风险控制的力度，尽可能地避免了网络欺诈行为的发生。

[0174] 参照图10所示，是本申请另一实施例中一种网络虚拟账户的风险控制方法的流程图。

[0175] 步骤810，分别获取第一虚拟账户的原始关联数据和第二虚拟账户的原始关联数据，所述原始关联数据包括至少一个关联维度的数据；

[0176] 在本申请实施例中，可以获取一个或多个关联维度的数据。

[0177] 步骤820，计算至少一个关联维度下第一虚拟账户与第二虚拟账户间的关联度，包括：基于数量因素、时间因素、级联因素其中至少一个，计算第一虚拟账户和第二虚拟账户间的关联度；

[0178] 本申请实施例对于每个关联维度，可以首先基于数量因素、时间因素、级联因素其中至少一个，计算两个虚拟账户之间的关联值。

[0179] 那么当只有一个关联维度时，该关联维度的关联值就是两个虚拟账户之间的关联度。当有多个关联维度时，该多个关联维度的关联值进行求和就得到两个虚拟账户之间的关联度

[0180] 步骤830，根据所述第一虚拟账户与第二虚拟账户之间的关联度，基于所述第一虚拟账户与第二虚拟账户中任意一个虚拟账户的风险状态，确定另一个虚拟账户的风险状态。

[0181] 本申请实施例的相关步骤的原理参照前述类似步骤，本申请实施例不对其加以限制。

[0182] 综上所述，在分析两个虚拟用户时，可以使用一个或多个关联维度的数据(如IP、cookie、机器指纹等等)进行分析。打破了传统分析方法的限制，能够更加准确地分析出网络虚拟用户的真实关联。进而，通过这种准确的分析方法，可以准确地识别出与恶意用户关联度高的其他用户也可能为恶意用户，大大提供了网络风险控制的力度，尽可能地避免了网络欺诈行为的发生。

[0183] 需要说明的是，对于前述的方法实施例，为了简单描述，故将其都表述为一系列的动作组合，但是本领域技术人员应该知悉，本申请并不受所描述的动作顺序的限制，因为依据本申请，某些步骤可以采用其他顺序或者同时进行。其次，本领域技术人员也应该知悉，说明书中所描述的实施例均属于优选实施例，所涉及的动作并不一定是本申请所必需的。

[0184] 基于上述方法实施例的说明，本申请还提供了相应的系统实施例。

[0185] 参照图11所示，是本申请实施例所述一种网络虚拟用户的风险控制系统结构图。

[0186] 所述系统可以包括以下模块：

[0187] 数据获取模块10，用于分别获取第一虚拟用户的原始关联数据和第二虚拟用户的原始关联数据，所述原始关联数据包括多维度的数据，每个维度的数据都能标识用户身份，

每个维度的数据作为一个关联维度；

[0188] 关联计算模块20，用于基于所述第一虚拟用户和第二虚拟用户的原始关联数据，分别计算每个关联维度下第一虚拟用户与第二虚拟用户间的关联值；

[0189] 关联汇总模块30，用于将各个关联维度下的关联值进行汇总，得到第一虚拟用户与第二虚拟用户间的真实关联度；

[0190] 风险判定模块40，用于判定所述第一虚拟用户与第二虚拟用户间的真实关联度是否超过预设阈值，如果超过，则当其中一个用户为恶意用户时，确定出与其关联的另一个用户也为恶意用户。

[0191] 在另一优选实施例中，所述关联计算模块20可以在每个关联维度下，都基于数量因素、时间因素和级联因素，计算第一虚拟用户和第二虚拟用户间的关联值。

[0192] 在另一优选实施例中，所述数量因素可采用求和函数。

[0193] 在另一优选实施例中，所述时间因素和级联因素都可采用倒数函数。

[0194] 在另一优选实施例中，所述关联计算模块20可以通过以下公式计算每个关联维度下第一虚拟用户和第二虚拟用户间的关联值：

[0195] $\Sigma_x \Sigma_{level} \Sigma_t (1/t)*(1/level)$；

[0196] 其中，Σ表示求和函数，x表示某个关联维度下的关联数量，level表示级联层次，t表示时间。

[0197] 在另一优选实施例中，所述多维度的原始关联数据可以包括：IP，和/ 或cookie，和/或机器指纹，和/或手机号码，和/或电话号码，和/或传真，和 /或电子邮箱，和/或地址，和/或登录账号。

[0198] 在另一优选实施例中，所述数据获取模块10可以包括以下子模块：

[0199] 第一获取子模块，用于获取第一虚拟用户的原始关联数据；

[0200] 第二获取子模块，用于基于所述第一虚拟用户的原始关联数据，查找与第一虚拟用户关联的第二虚拟用户，以及所述第二虚拟用户的原始关联数据。

[0201] 在另一优选实施例中，所述第二获取子模块可以包括以下子单元：

[0202] 第一查询子单元，用于对每个关联维度，根据第一虚拟用户的标识查找该用户使用的维度数据；

[0203] 第二查询子单元，用于利用所述查找到的维度数据，继续查找其中每个维度数据对应的用户列表；

[0204] 去重整理子单元，用于将每个关联维度下查找到的用户列表进行去重整理，最后得到的用户列表中标识的所有用户即为与所述第一虚拟用户关联的第二虚拟用户。

[0205] 在另一优选实施例中，所述关联汇总模块30可以包括以下子模块：

[0206] 求和子模块，用于将各个关联维度下的关联值进行求和。

[0207] 对于上述系统实施例而言，由于其与方法实施例基本相似，所以描述的比较简单，相关之处参见方法实施例的部分说明即可。

[0208] 上述风险控制系统可以准确地分析出网络虚拟用户间的真实关联情况，进而可以准确地识别出与恶意用户关联度高的其他用户也可能为恶意用户，大大提供了网络风险控制的力度，尽可能地避免了网络欺诈行为的发生。

[0209] 本说明书中的各个实施例均采用递进的方式描述，每个实施例重点说明的都是与

其他实施例的不同之处，各个实施例之间相同相似的部分互相参见即可。

[0210] 需要说明的是，在本文中，诸如第一和第二等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来，而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。

[0211] 而且，上文中的“和/或”表示本文既包含了“和”的关系，也包含了“或”的关系，其中：如果方案A与方案B是“和”的关系，则表示某实施例中可以同时包括方案A和方案B；如果方案A与方案B是“或”的关系，则表示某实施例中可以单独包括方案A，或者单独包括方案B。

[0212] 以上对本申请所提供的一种网络虚拟用户的风险控制方法及系统，进行了详细介绍，本文中应用了具体个例对本申请的原理及实施方式进行了阐述，以上实施例的说明只是用于帮助理解本申请的方法及其核心思想；同时，对于本领域的一般技术人员，依据本申请的思想，在具体实施方式及应用范围上均会有改变之处，综上所述，本说明书内容不应理解为对本申请的限制。

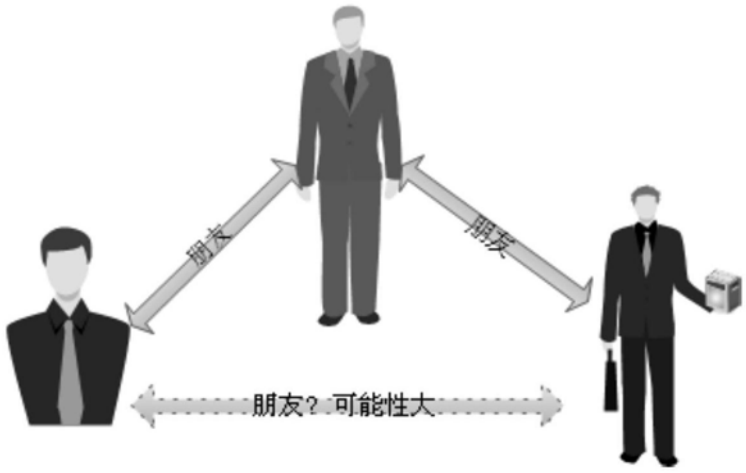


图1

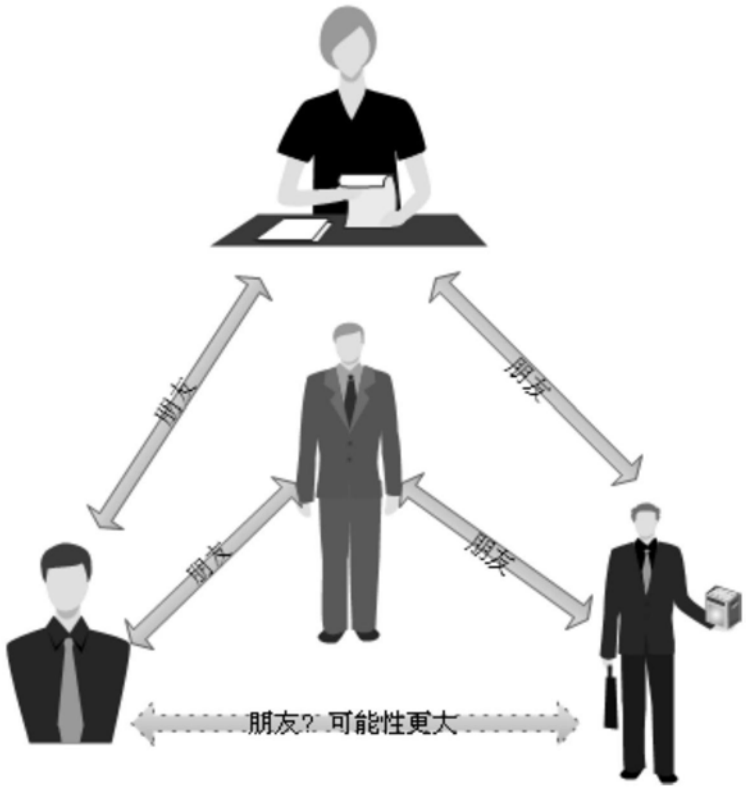


图2

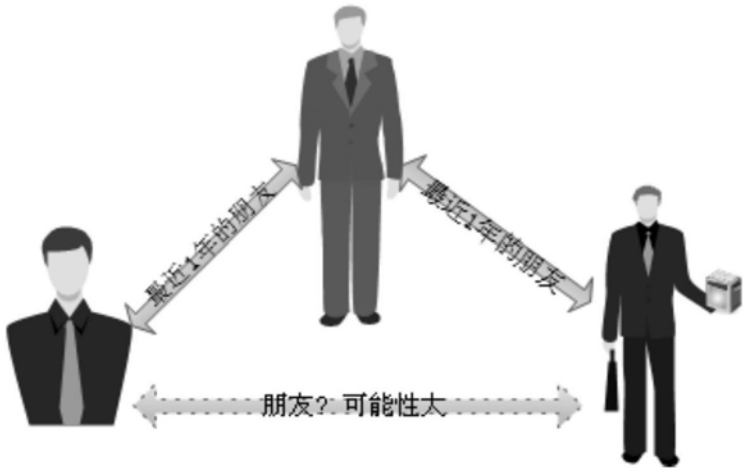

最近1年的朋友
最近1年的朋友
朋友？可能性大

图3.1

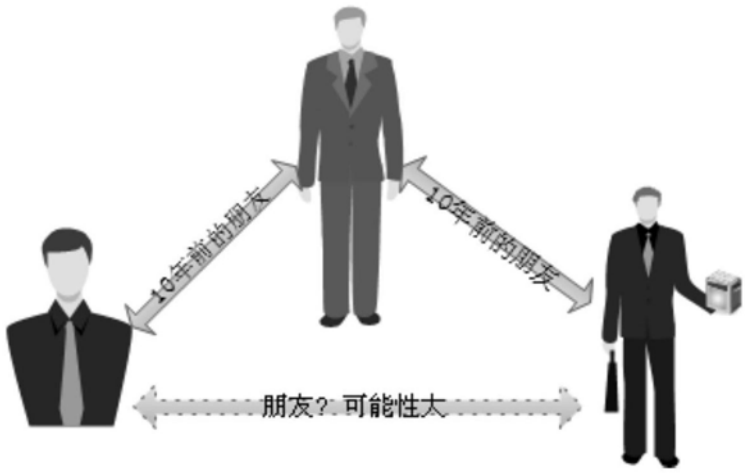

10年前的朋友
10年前的朋友
朋友？可能性大

图3.2

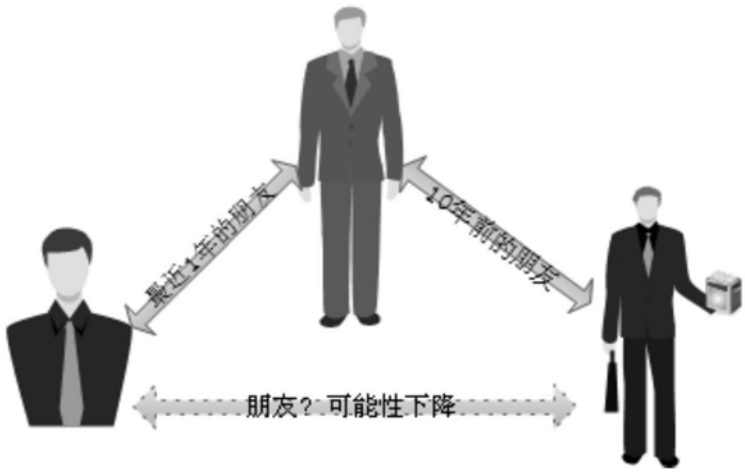

最近1年的朋友
10年前的朋友
朋友？可能性下降

图3.3

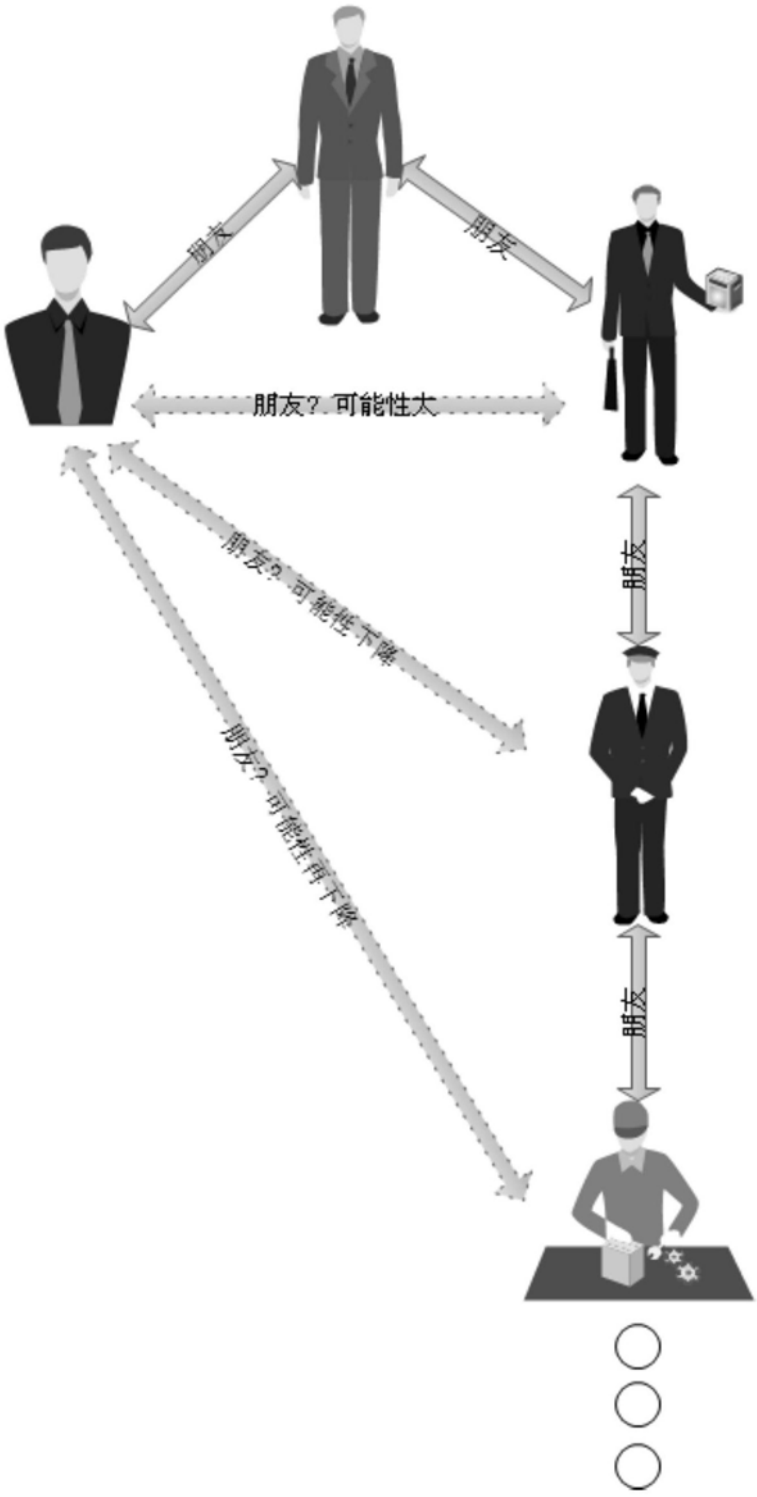


图4

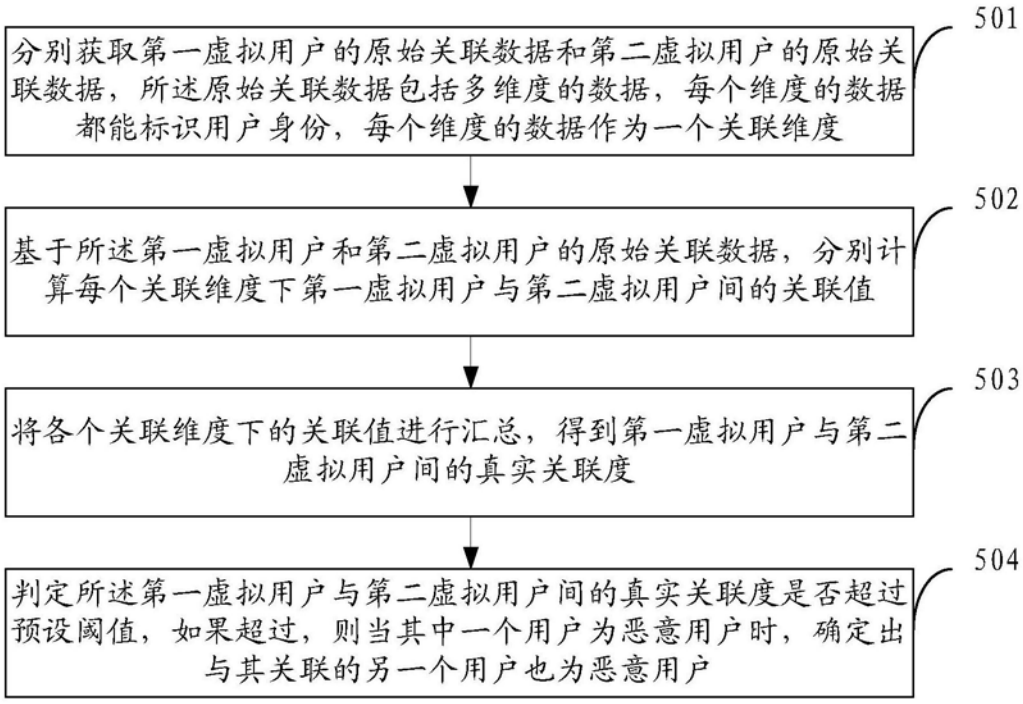

501
分别获取第一虚拟用户的原始关联数据和第二虚拟用户的原始关联数据，所述原始关联数据包括多维度的数据，每个维度的数据都能标识用户身份，每个维度的数据作为一个关联维度
502
基于所述第一虚拟用户和第二虚拟用户的原始关联数据，分别计算每个关联维度下第一虚拟用户与第二虚拟用户间的关联值
503
将各个关联维度下的关联值进行汇总，得到第一虚拟用户与第二虚拟用户间的真实关联度
504
判定所述第一虚拟用户与第二虚拟用户间的真实关联度是否超过预设阈值，如果超过，则当其中一个用户为恶意用户时，确定出与其关联的另一个用户也为恶意用户

图5

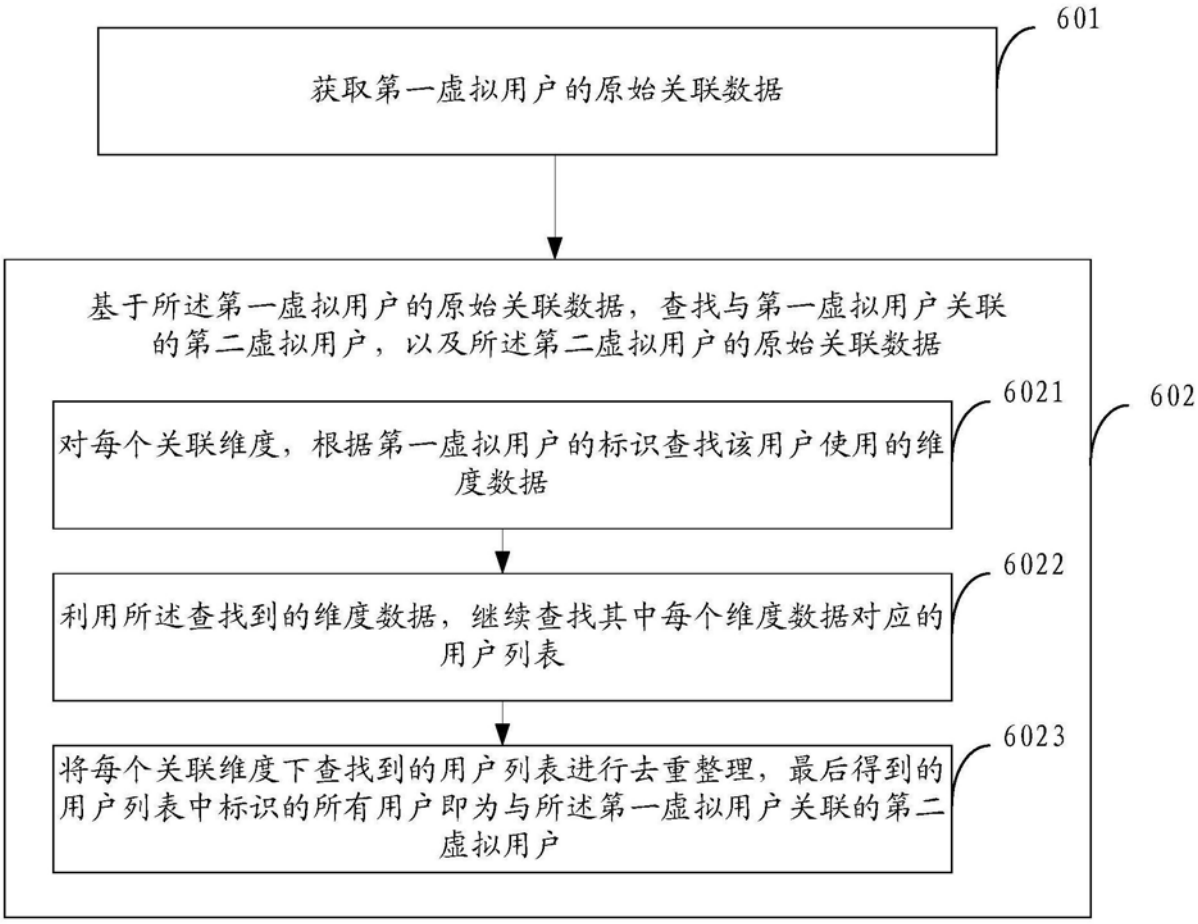

601
获取第一虚拟用户的原始关联数据
基于所述第一虚拟用户的原始关联数据，查找与第一虚拟用户关联的第二虚拟用户，以及所述第二虚拟用户的原始关联数据
602
6021
对每个关联维度，根据第一虚拟用户的标识查找该用户使用的维度数据
6022
利用所述查找到的维度数据，继续查找其中每个维度数据对应的用户列表
6023
将每个关联维度下查找到的用户列表进行去重整理，最后得到的用户列表中标识的所有用户即为与所述第一虚拟用户关联的第二虚拟用户

图6


IP
122.25.36.***
Cookie
491233421.***2
机器指纹
jflewfoe/v34;h0
电话号码
010-8588****
使用过
使用过
同一人、同伙、互相...？可能性大

图7

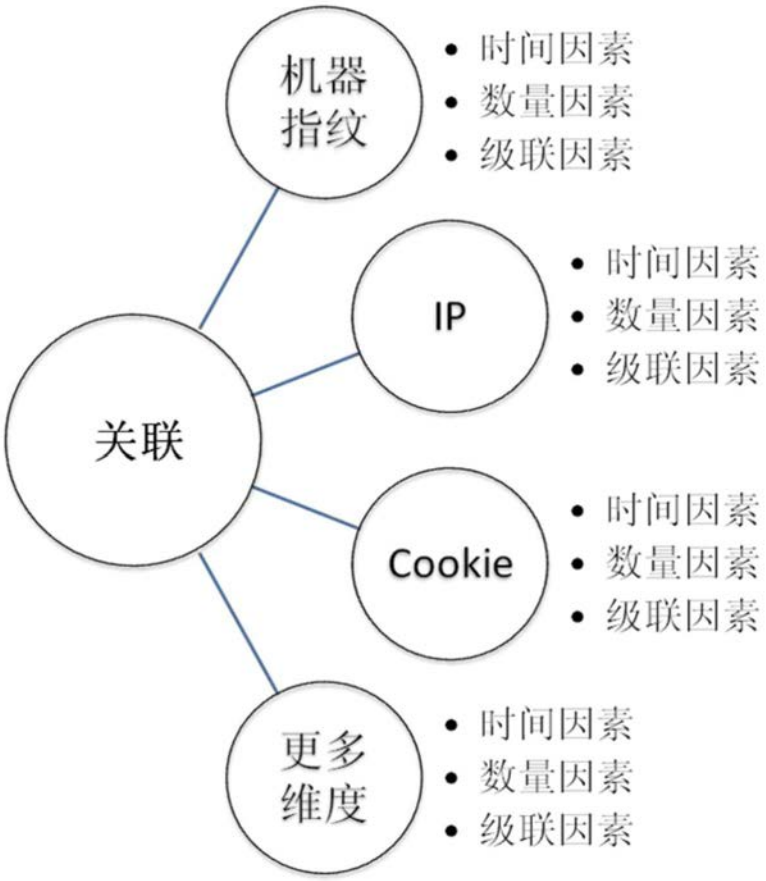

关联
机器指纹
• 时间因素
• 数量因素
• 级联因素
IP
• 时间因素
• 数量因素
• 级联因素
Cookie
• 时间因素
• 数量因素
• 级联因素
更多维度
• 时间因素
• 数量因素
• 级联因素

图8

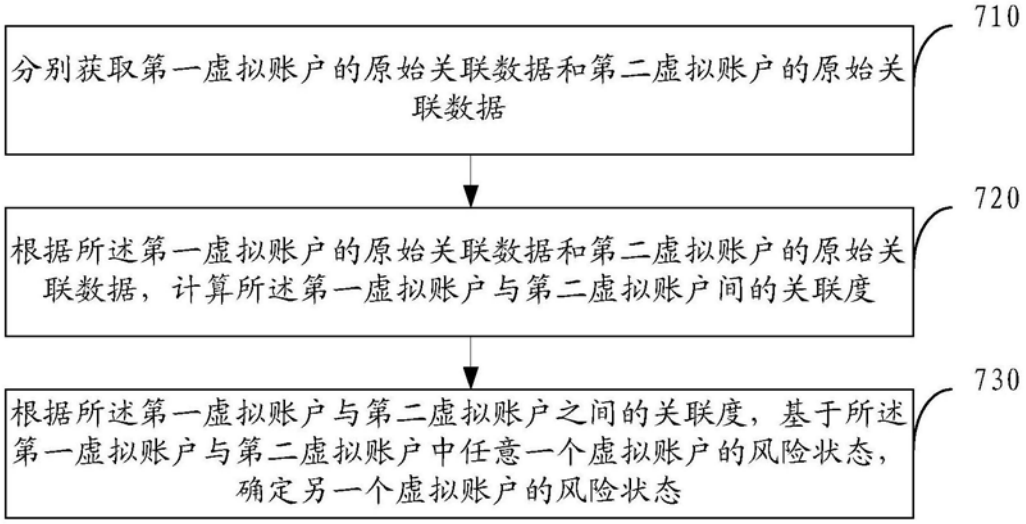

710
分别获取第一虚拟账户的原始关联数据和第二虚拟账户的原始关联数据
720
根据所述第一虚拟账户的原始关联数据和第二虚拟账户的原始关联数据，计算所述第一虚拟账户与第二虚拟账户间的关联度
730
根据所述第一虚拟账户与第二虚拟账户之间的关联度，基于所述第一虚拟账户与第二虚拟账户中任意一个虚拟账户的风险状态，确定另一个虚拟账户的风险状态

图9

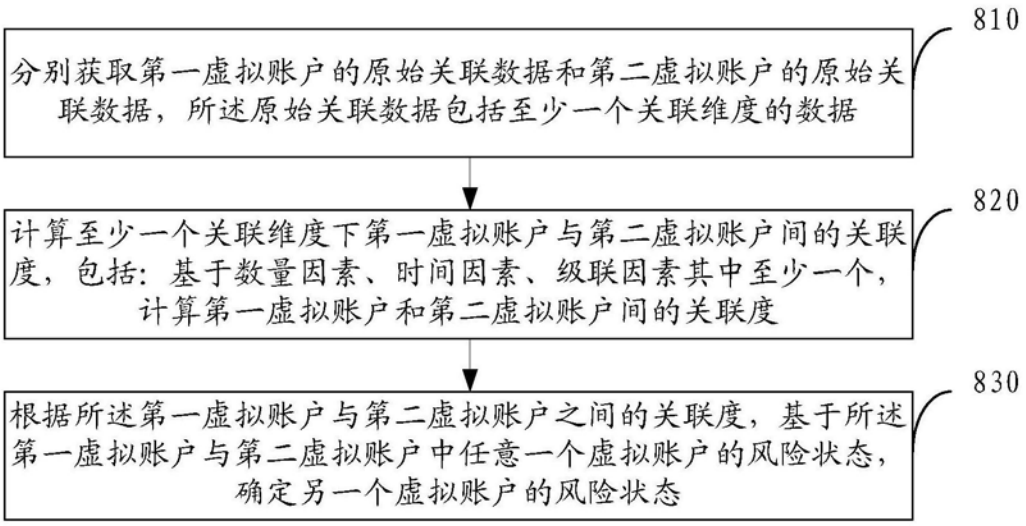

810
分别获取第一虚拟账户的原始关联数据和第二虚拟账户的原始关联数据，所述原始关联数据包括至少一个关联维度的数据
820
计算至少一个关联维度下第一虚拟账户与第二虚拟账户间的关联度，包括：基于数量因素、时间因素、级联因素其中至少一个，计算第一虚拟账户和第二虚拟账户间的关联度
830
根据所述第一虚拟账户与第二虚拟账户之间的关联度，基于所述第一虚拟账户与第二虚拟账户中任意一个虚拟账户的风险状态，确定另一个虚拟账户的风险状态

图10


数据获取模块10
关联计算模块20
关联汇总模块30
风险判定模块40

图11