

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

H04L 12/24 (2006.01)

G06F 17/30 (2006.01)

H04L 12/58 (2006.01)



[12] 发明专利说明书

专利号 ZL 200610104256.2

[45] 授权公告日 2009年11月4日

[11] 授权公告号 CN 100558045C

[22] 申请日 2006.8.7

[21] 申请号 200610104256.2

[73] 专利权人 华为技术有限公司

地址 518129 广东省深圳市龙岗区坂田华为总部办公楼

[72] 发明人 闵国兵 钟杰萍 汪芳山 方琦

[56] 参考文献

CN1771712A 2006.5.10

US2002173971A1 2002.11.21

US20060106876A1 2006.5.18

US20060005113A1 2006.1.5

US5794050A 1998.8.11

审查员 李倩

[74] 专利代理机构 北京德琦知识产权代理有限公司

代理人 王琦 王诚华

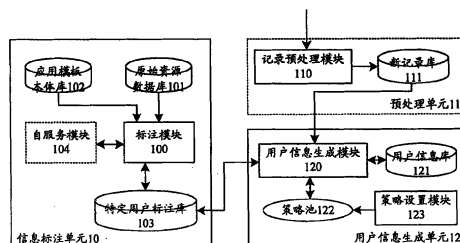
权利要求书4页 说明书25页 附图3页

[54] 发明名称

一种生成通信用户描述信息的系统及方法

[57] 摘要

本发明公开了一种生成通信用户描述信息的系统，该系统包括用于对特定通信用户的原始信息进行标注的信息标注单元(10)，以及用于根据通信记录生成通信用户描述信息的用户信息生成单元(12)；本发明同时还公开了一种生成通信用户描述信息的方法，通过将本体论应用到通信领域中，通过对用户的真实通信记录的分析，获得了用户需求信息并生成了通信用户描述信息。而且，生成的通信用户描述信息，真实地反应了通信用户的特征与需求，保证了生成的通信用户描述信息的准确性。本发明方法适用于用户的语音通信、数据通信、短信等通信行为。



1. 一种生成通信用户描述信息的系统，其特征在于，该系统包括：用于对特定通信用户的原始信息进行标注的信息标注单元（10），以及用于生成通信用户描述信息的用户信息生成单元（12）；

其中，信息标注单元（10）包括标注模块（100）、原始资源数据库（101）、应用模板本体库（102）和特定用户标注库（103）；

原始资源数据库（101），用于记录需要标注的用户的原始信息；

应用模板本体库（102），用于针对不同的应用，建立描述该应用的类与类属性的本体的应用模板并存储；

标注模块（100），从原始资源数据库（101）中获取需要标注的用户的原始信息，根据该原始信息所属类，调用应用模板本体库（102）中相应类的应用模板，将原始信息中对应应用模板的信息值填入该应用模板中，实现对该原始信息的标注，并将经过标注的原始信息存储在特定用户标注库（103）中；

用户信息生成单元（12）包括用户信息生成模块（120）、用户信息库（121）、策略池（122）和策略设置模块（123）；

用户信息生成模块（120），从外界获取通信记录信息，查询特定用户标注库（103），并判断特定用户标注库（103）中是否存储有通信记录中的联系方的标识字段值，若未存储有，则按照策略池（122）中的非特定用户的信息传递策略，将满足条件的联系方标识信息添加到通信主体的通信用户描述信息中并存储在用户信息库（121）中；若存储有，则按照策略池（122）中的特定用户的信息传递策略，将满足条件的联系方标识信息，以及该联系方标识信息对应的存储在特定用户标注库（103）中的用户标注信息，添加到通信主体的通信用户描述信息中并存储在用户信息库（121）中；

策略设置模块（123），用于设置各种用户信息传递策略，并存储到策略池（122）中；

策略池（122），用于存储各种用户信息传递策略，并提供给用户信息生

成模块(120)使用;

用户信息库(121),用于存储通信主体的通信用户描述信息文件。

2. 根据权利要求1所述的系统,其特征在于,所述系统还包括:用于对通信记录进行预处理的预处理单元(11),该预处理单元包括记录预处理模块(110)和新记录库(111),其中,

记录预处理模块(110),用于对从外界获取的通信记录信息进行预处理,从原始通信记录字段中提取预先设定的目标字段组成新记录并存储,同时所述新记录还包括按照预设要求对通信记录进行处理后产生的附加字段;

新记录库(111),用于存储由目标字段和附加字段组成的新的通信记录,并提供给用户信息生成模块(120)使用。

3. 根据权利要求1或2所述的系统,其特征在于,所述信息标注单元(10)还包括:自服务模块(104),用于提供一个面向用户的接口,使用自服务模块(104),并通过所述标注模块(100)对特定用户标注库(103)中标注的结果进行更新。

4. 根据权利要求3所述的系统,其特征在于,所述自服务模块(104)为一操作平台,或者客户端/服务器C/S架构中的客户端。

5. 根据权利要求1所述的系统,其特征在于,所述标注模块(100)中进一步包括所述应用模板中的字段与其对应的模糊字段的映射表;

所述标注模块(100)还包括:对所述应用模板中的字段对应的模糊字段进行分析以提取关键字段的切词功能。

6. 根据权利要求1所述的系统,其特征在于,所述应用模板包括应用模板名称,及描述该应用模板名称所指示的类的类属性的应用模板参数;所述应用模板名称与所述类的名称相同。

7. 根据权利要求6所述的系统,其特征在于,所述应用模板还包括:不同类之间的关联关系。

8. 一种生成通信用户描述信息的方法,其特征在于,预先设置决定是否将联系方信息加入通信主体的通信用户描述信息中的用户信息传递策略,该方

法还包括以下步骤:

A. 获取通信记录信息,判断通信记录信息中的联系方标识信息是否存在于生成的用户标注信息中,若存在,则进入步骤 C; 否则,进入步骤 B; 所述生成用户标注信息的方法包括: 预先设置针对不同的应用,用于描述该应用的类与类属性的本体的应用模板;

A11. 提取一条原始资源数据记录,并获取该原始资源数据记录的所属类的类名称;

A12. 根据获得的类名称,调用对应该类名称的应用模板,并根据调用的应用模板中包括的字段,将该原始资源数据记录中对应的字段值记录在该应用模板中;

A13. 存储完成记录的应用模板;

B. 按照预设的用户信息传递策略中的非特定用户的信息传递策略,将满足条件的通信记录信息中的联系方标识信息记录到通信主体的通信用户描述信息中,之后结束本流程;

C. 按照预设的用户信息传递策略中的特定用户的信息传递策略,将满足条件的通信记录信息中的联系方标识信息,以及该联系方标识信息对应的用户标注信息记录到通信主体的通信用户描述信息中。

9. 根据权利要求 8 所述的方法,其特征在于,若所述应用模板中的部分字段在原始资源数据记录中不存在,则步骤 A12 中所述将原始资源数据记录中对应的字段值记录在应用模板中的方法为: 将应用模板中该字段的值设置为空。

10. 根据权利要求 8 所述的方法,其特征在于,若所述应用模板中的部分字段在原始资源数据记录中不存在,则步骤 A12 中所述将原始资源数据记录中对应的字段值记录在应用模板中的方法为:

预先设置应用模板中的字段与其对应的模糊字段的映射表;

查询映射表,判断所述原始资源数据记录中是否存在所述应用模板中该字段的模糊字段,若存在所述应用模板中该字段的模糊字段,从该模糊字段中分析提取关键字段,并将该关键字段的值记录在所述应用模板的对应字段中; 若不存在所述应用模板中该字段的模糊字段,将所述应用模板中该字段的值设置为空。

11. 根据权利要求 10 所述的方法, 其特征在于, 所述分析提取关键字段的方法为: 切词分析。

12. 根据权利要求 8 所述的方法, 其特征在于, 若获取的通信记录信息中不存在所述用户信息传递策略的条件中描述的字段, 则步骤 A 中所述判断之前, 还包括:

对所述获取的通信记录信息进行预处理, 从原始通信记录字段中提取预先设定的目标字段, 以及按照预设要求对所述通信记录信息进行处理并产生的附加字段, 组成新的通信记录信息并存储。

13. 根据权利要求 8 所述的方法, 其特征在于, 步骤 C 中所述记录用户标注信息的方法为:

遍历所述通信主体用户的通信用户描述信息中预设的偏好属性, 判断该偏好属性中是否存在所述用户标注信息所属的类, 若不存在所述用户标注信息所属的类, 则在该通信主体用户的偏好属性中添加所述用户标注信息所属的类及其中的可传递参数;

若存在所述用户标注信息所属的类, 则进一步判断该偏好属性中是否存在所述用户标注信息中的可传递参数, 若存在所述用户标注信息中的可传递参数, 结束; 若不存在所述用户标注信息中的可传递参数, 则将所述用户标注信息中的可传递参数添加在该偏好属性中。

14. 根据权利要求 8 所述的方法, 其特征在于, 该方法进一步包括: 在所述通信用户描述信息中, 根据预设的指定字段的组合运算关系, 建立通信主体与所述非特定用户的联系频繁度或亲密度, 或者建立通信主体与所述特定用户的联系频繁度或亲密度。

15. 根据权利要求 8 所述的方法, 其特征在于, 所述应用模板是基于本体语言创建的。

16. 根据权利要求 15 所述的方法, 其特征在于, 所述应用模板包括应用模板名称, 及描述该应用模板名称所指示的类的相关类属性的应用模板参数。

一种生成通信用户描述信息的系统及方法

技术领域

本发明涉及通信信息收集技术，尤指一种生成通信用户描述信息的系统及方法。

背景技术

随着现代通信技术的发展，人与人之间的交流工具日趋多样化，从以语音为代表的固定电话和移动电话，以文字信息为主的 Email、移动 Email，到融合了语音、文字甚至图片的即时消息（IM）交流；从环球网（Web）1.0 的论坛到 Web2.0 的博客（Blog）。同时各种交流方式也渐趋融合，比如，用户可以仅仅通过一个移动终端，即可完成话音通信、Email 收发、IM 交流，更可通过各种数据业务如移动 Blog、图片 Blog、播客（Podcast）等完成信息交流。

有关资料表明，在现有条件下，一个具有 70 万 G 网用户的地区，一个月将产生 1.7 亿条话单记录。人与人之间的通信行为可以建立起一个巨大的社会关系网，同时，在这些交流中蕴涵着丰富的反应用户特征与需求的通信用户描述信息，比如，某人经常打证券公司的服务电话，即可推断此人有可能是股民，投资理财类的产品可能会引起该用户的兴趣，由此，可以将证券公司的某些属性传递给该用户。在以用户为中心的今天，如何获取通信用户描述信息成为一个很有意义的课题。

另一方面，本体论已然成为国内外研究热点，在多个领域出现了具体应用，其研究集中在知识工程、本体工程、信息组织与检索和语义 Web 等方面。各种本体描述语言和本体创建工具也逐渐被开发和应用。比如，Protégé 是由斯坦福大学医学院医学信息中心开发的免费、开放源码的本体编辑器和知

识管理平台。Jena是惠普(HP)公司开发的面向语义网的本体解析、规则推理等应用的开发包,包含的内容比较全面。本体语言方面,Web本体语言OWL已经成为W3C的推荐标准。在传统行业以及新兴行业应用本体论已经成为一种趋势。

将本体论应用在通信世界中,用本体语言来描述各通信主体的属性特征,进而捕捉属性特征在社会网络中的传递情况,这对于获取作为主要通信主体的用户的描述信息具有重要意义。

Google公司的Gmail邮箱服务在推出1G邮箱为用户服务的同时,通过网友之间的相互推荐建立用户关系网,获取了用户的一些信息比如社会关系情况,从而生成了一些通信用户描述信息。Gmail的新用户必须获得老用户的邀请才能取得邮箱帐号,通过网友之间邀请权限的使用,收集用户的社会关系属性,建立一个社会网络。然而,真正的实施情况是Gmail的老用户邀请的并不限于他的朋友,在Gmail推出初期,很多Gmail拥有者在网上公开销售邀请权限,而当Gmail普及之后,可以很容易在网上获取Gmail邀请权限。这样,导致了基于邀请权限建立起来的社会关系并不准确,也就是生成的通信用户描述信息不准确。

另一种生成通信用户描述信息的方法是:通过从服务器侧分析用户的邮件收发记录获取用户与哪些类型的邮箱地址通信,从而根据这些邮箱地址信息生成用户描述信息,但是邮件提供商通过分析邮件的收发记录只能建立简单的联系网络,而这些很少的通信用户描述信息是不能真实反应用户特征与需求的,是不准确的。

发明内容

有鉴于此,本发明的主要目的在于提供一种生成通信用户描述信息的系统,该系统能够自动生成通信用户描述信息,保证生成的通信用户描述信息的准确性。

本发明的另一目的在于提供一种生成通信用户描述信息的方法,该方法

能够自动生成通信用户描述信息，保证生成的通信用户描述信息的准确性。

为达到上述目的，本发明的技术方案具体是这样实现的：

一种生成通信用户描述信息的系统，该系统包括：用于对特定通信用户的原始信息进行标注的信息标注单元，以及用于生成通信用户描述信息的用户信息生成单元；

其中，信息标注单元包括标注模块、原始资源数据库、应用模板本体库和特定用户标注库；

原始资源数据库，用于记录需要标注的用户的原始信息；

应用模板本体库，用于针对不同的应用，建立描述该应用的类与类属性的本体的应用模板并存储；

标注模块，从原始资源数据库中获取需要标注的用户的原始信息，根据该原始信息所属类，调用应用模板本体库中相应类的应用模板，将原始信息中对应应用模板的信息值填入该应用模板中，实现对该原始信息的标注，并将经过标注的原始信息存储在特定用户标注库中；

用户信息生成单元包括用户信息生成模块、用户信息库、策略池和策略设置模块；

用户信息生成模块，从外界获取通信记录信息，查询特定用户标注库，并判断特定用户标注库中是否存储有通信记录中的联系方的标识字段值，若未存储有，则按照策略池中的非特定用户的信息传递策略，将满足条件的联系方标识信息添加到通信主体的通信用户描述信息中并存储在用户信息库中；若存储有，则按照策略池中的特定用户的信息传递策略，将满足条件的联系方标识信息，以及该联系方标识信息对应的存储在特定用户标注库中的用户标注信息，添加到通信主体的通信用户描述信息中并存储在用户信息库中；

策略设置模块，用于设置各种用户信息传递策略，并存储到策略池中；

策略池，用于存储各种用户信息传递策略，并提供给用户信息生成模块使用；

用户信息库，用于存储通信主体的通信用户描述信息文件。

所述系统还包括：用于对通信记录进行预处理的预处理单元，该预处理单元包括记录预处理模块和新记录库，其中，

记录预处理模块，用于对从外界获取的通信记录信息进行预处理，从原始通信记录字段中提取预先设定的目标字段组成新记录并存储，同时所述新记录还包括按照预设要求对通信记录进行处理后产生的附加字段；

新记录库，用于存储由目标字段和附加字段组成的新的通信记录，并提供给用户信息生成模块使用。

所述信息标注单元还包括：自服务模块，用于提供一个面向用户的接口，使用自服务模块，并通过所述标注模块对特定用户标注库中标注的结果进行更新。

所述自服务模块为一操作平台，或者客户端/服务器 C/S 架构中的客户端。

所述标注模块中进一步包括所述应用模板中的字段与其对应的模糊字段的映射表；

所述标注模块还包括：对所述应用模板中的字段对应的模糊字段进行分析以提取关键字段的切词功能。

所述应用模板包括应用模板名称，及描述该应用模板名称所指示的类的相关类属性的应用模板参数；所述应用模板名称与所述类的名称相同。

所述应用模板还包括：不同类之间的关联关系。

一种生成通信用户描述信息的方法，预先设置决定是否将联系方信息加入通信主体的通信用户描述信息中的用户信息传递策略，该方法还包括以下步骤：

A. 获取通信记录信息，判断通信记录信息中的联系方标识信息是否存在于生成的用户标注信息中，若存在，则进入步骤 C；否则，进入步骤 B；所述生成用户标注信息的方法包括：预先设置针对不同的应用，用于描述该应用的类与类属性的本体的应用模板；

A11. 提取一条原始资源数据记录，并获取该原始资源数据记录的所属类的类名称；

A12. 根据获得的类名称，调用对应该类名称的应用模板，并根据调用的应用模板中包括的字段，将该原始资源数据记录中对应的字段值记录在该应用模板中；

A13. 存储完成记录的应用模板；

B. 按照预设的用户信息传递策略中的非特定用户的信息传递策略，将满足

条件的通信记录信息中的联系方标识信息记录到通信主体的通信用户描述信息中，之后结束本流程；

C. 按照预设的用户信息传递策略中的特定用户的信息传递策略，将满足条件的通信记录信息中的联系方标识信息，以及该联系方标识信息对应的用户标注信息记录到通信主体的通信用户描述信息中。

若所述应用模板中的部分字段在原始资源数据记录中不存在，则步骤 A12 中所述将原始资源数据记录中对应的字段值记录在应用模板中的方法为：将应用模板中该字段的值设置为空。

若所述应用模板中的部分字段在原始资源数据记录中不存在，则步骤 A12 中所述将原始资源数据记录中对应的字段值记录在应用模板中的方法为：

预先设置应用模板中的字段与其对应的模糊字段的映射表；

查询映射表，判断所述原始资源数据记录中是否存在所述应用模板中该字段的模糊字段，若存在所述应用模板中该字段的模糊字段，从该模糊字段中分析提取关键字段，并将该关键字段的值记录在所述应用模板的对应字段中；若不存在所述应用模板中该字段的模糊字段，将所述应用模板中该字段的值设置为空。

所述分析提取关键字段的方法为：切词分析。

若获取的通信记录信息中不存在所述用户信息传递策略的条件中描述的字段，则步骤 A 中所述判断之前，还包括：

对所述获取的通信记录信息进行预处理，从原始通信记录字段中提取预先设定的目标字段，以及按照预设要求对所述通信记录信息进行处理并产生的附加字段，组成新的通信记录信息并存储。

步骤 C 中所述记录用户标注信息的方法为：

遍历所述通信主体用户的通信用户描述信息中预设的偏好属性，判断该偏好属性中是否存在所述用户标注信息所属的类，若不存在所述用户标注信息所属的类，则在该通信主体用户的偏好属性中添加所述用户标注信息所属的类及其中的可传递参数；

若存在所述用户标注信息所属的类，则进一步判断该偏好属性中是否存

在所述用户标注信息中的可传递参数，若存在所述用户标注信息中的可传递参数，结束；若不存在所述用户标注信息中的可传递参数，则将所述用户标注信息中的可传递参数添加在该偏好属性中。

该方法进一步包括：在所述通信用户描述信息中，根据预设的指定字段的组合运算关系，建立通信主体与所述非特定用户的联系频繁度或亲密度，或者建立通信主体与所述特定用户的联系频繁度或亲密度。

所述应用模板是基于本体语言创建的。

所述应用模板包括应用模板名称，及描述该应用模板名称所指示的类的相关类属性的应用模板参数。

由上述技术方案可见，本发明通过将本体论应用到通信领域中，通过对用户的真实通信记录的分析，获得了用户需求信息并生成了通信用户描述信息。而且，生成的通信用户描述信息，真实地反应了通信用户的特征与需求，保证了生成的通信用户描述信息的准确性。

本发明方法适用于用户的语音通信、数据通信、短信等通信行为。

附图说明

图 1 是本发明提供的实施例中生成通信用户描述信息的系统组成结构示意图；

图 2 是本发明提供的实施例中生成通信用户描述信息的方法的流程图；

图 3 是本发明提供的实施例中实现信息标注的流程图；

图 4 是本发明提供的实施例中一种记录用户应用模板的方法的流程图；

图 5 是本发明实施例的流程图。

具体实施方式

本发明的核心思想是：预先设置用户信息传递策略；获取通信记录信息，判断通信记录信息中的联系方标识信息是否存在于生成的用户标注信息中，若不存在，按照预设的用户信息传递策略中的非特定用户的信息传递策略，

将满足条件的联系方标识信息记录到通信主体的通信用户描述信息中；若存在，按照预设的用户信息传递策略中的特定用户的信息传递策略，将满足条件的联系方标识信息，以及该联系方标识信息对应的用户标注信息记录到通信主体的通信用户描述信息中。

为使本发明的目的、技术方案及优点更加清楚明白，以下参照附图并举较佳实施例，对本发明进一步详细说明。

本发明方法适用于用户的语音通信、数据通信、短信等通信行为。本发明将以语音通信作为实施例来描述实现过程。

图 1 是本发明提供的实施例中生成通信用户描述信息的系统组成结构示意图，本发明系统包括用于对特定通信用户的原始信息进行标注的信息标注单元 10，以及用于根据通信记录生成通信用户描述信息的用户信息生成单元 12，以下将通信用户简称为用户，下面详细介绍各单元的组成及连接关系。

(一) 信息标注单元 10 包括标注模块 100、原始资源数据库 101、应用模板本体库 102 和特定用户标注库 103，其中，

原始资源数据库 101，用于记录需要标注的用户的原始信息，包括描述用户的标识信息和属性信息，比如企业原始的黄页注册数据：企业所属类、企业名称、电话号码、业务介绍、所在地区、地址、传真、邮编、地址、Email 等。其中，所属类记录该公司所属的行业目录叶节点，公司电话号码作为标识信息，剩余参数作为属性描述信息。需要说明的是，本文中提到的特定用户是指经过标注的通信用户。

假设在 IM 通信网络中，对于 IM 号也区分企业用户和普通用户，那么原始资源数据库指 IM 的企业用户信息记录。同样的处理可以是 Email 等各种具有号码的服务中。

应用模板本体库 102，用于针对不同的应用，参照原始资源数据的分类方式，建立各大类及其子类、以及各类属性的本体，定义类与类之间的关系和类属性与类属性之间的关联。同时，定义类与类属性之间的关系，指明一

个类及其子类应该由哪些类属性来描述，最终由这个关系导出类的应用模板。需要说明的是，如果原始资源数据没有进行分类，可以人工为原始资源数据分类，同时创建类库的本体。这里，应用模板是基于本体语言创建的，创建的方法属于现有技术，可以通过目前许多本体创建的工具来实现，如背景技术中提到的 Protégé 平台等，这里不再详述。

一个应用模板由两部分组成：应用模板名称，用于标识该应用模板是描述哪个类的，应用模板名称与类的名称一致；应用模板参数，描述该应用模板名称所指示的类的相关类属性，在设置应用模板参数时，每个应用模板参数对应一个类型标签。应用模板参数的类型可以分为传递参数、非传递参数，甚至更多类型，当一个参数被设为传递参数时，该参数可以传递给用户作为用户描述信息记录。

由于每个行业具有不同的特点，需要不同的应用模板参数对其属性进行描述，比如汽车类行业可能具有品牌类的属性，用于标识该公司经营或制造的汽车品牌，而餐饮类行业则可能不存在品牌类的属性，更多描述的是口味、装修风格等参数。所以，可以针对不同行业分类目录的叶节点，设置不同的应用模板。每个应用模板具有唯一分配的模板名称和针对该叶节点所在行业的应用模板参数描述。

下面以汽车销售的应用模板为例，说明汽车销售类特定用户可以具有的应用模板参数，如表 1 所示：

应用模板名称：汽车销售	
所含应用模板参数	
组 1：基本属性（固有属性）	
公司名称：	电话号码：
地址：	城市：
邮编：	网址：
组 2：扩展属性（服务属性）	
品牌：如宝马、富康、大众等	
服务项：如汽车销售、汽车保险、汽车保养等	

表 1

表 1 用文字的方式描述了一个汽车销售应用模板可能包含的内容，具体

实现中的参数根据实际情况而定，包括但不限于表 2 所示。

在实际的实现过程中，可以利用本体描述语言如环球网本体描述语言（OWL，Web Ontology Language）对应用模板进行描述和记录，具体如何进行描述属于现有技术，本发明强调的是针对不同的应用（类）建立不同的应用模板。

应用模板不仅仅为类标注各种应用模板参数，还可以建立不同类之间的关联关系，如汽车类中的汽车销售可以与化工类中的汽油销售建立关联关系，表示汽车销售的用户同时有可能是汽油销售的用户，这种传递关系能够丰富所收集的用户属性。

标注模块 100，是信息标注单元 10 的逻辑处理模块，用于对用户的原始信息进行标注。从原始资源数据库 101 中获取记录的需要标注的用户的原始信息，根据该原始信息所属类，调用应用模板本体库 102 中相应类的应用模板，将原始信息中对应应用模板的应用模板名称和应用模板参数项的参数值填入该应用模板中，从而获得特定用户信息，并将获得的特定用户信息存储在特定用户标注库 103 中。

对特定用户属性进行标注的过程中，可以针对不同类型行业的用户，调用不同的应用模板进行特定用户信息的标注。

特定用户标注库 103，用于存储经过标注的特定用户信息。

经过标注的特定用户信息包括但不限于特定用户的标识信息如电话号码、特定用户所属的应用模板名称、特定用户的描述信息。其中，特定用户的描述信息包括但不限于基本属性和扩展属性。特定用户信息在特定用户标注库 103 中以一种可行的存储方式进行存储，表 2 是存储的特定用户信息的一个实施例：

<特定用户 1>
<标识号码></标识号码>
<模板名称>汽车销售</模板名称>
<属性参数>
<组 1: 基本属性>

<企业名称></企业名称> <地区></地区> <地址></地址> <邮编 ></邮编 > <网址 ></网址 >
</组 1: 基本属性>
<组 2: 扩展属性 >
<品牌 ></品牌 > <服务项 ></服务项 >
</组 2: 扩展属性 >
</属性参数>
</特定用户 1 >

表 2

信息标注单元 10 还可以进一步包括自服务模块 104，用于提供一个面向用户的接口，让用户可以使用自服务模块 104，并通过标注模块 100 对特定用户标注库 103 中标注的结果进行更新。如果是面向的用户是系统管理员，那么这个相当于是一个操作平台如系统界面；如果面向的是被标注的用户，则采用客户端/服务器（C/S）的架构，此时，自服务模块 104 相当于客户端，标注模块 100 和特定用户标注库 103 相当于服务器。被标注的用户通过客户端登录到服务器，可以通过客户端进行注册、登录、信息更新等操作修改自身的标注的信息。这里，由于 C/S 结构是比较常用的现有架构，且属于本领域技术人员公知技术，所以这里不再详述。

（二）用户信息生成单元 12 包括用户信息生成模块 120、用户信息库 121、策略池 122 和策略设置模块 123。其中，

用户信息生成模块 120，是本发明系统的一个核心模块，用于根据从策略池 122 中查询并调用的用户信息传递策略，决定是否和如何生成通信用户描述信息，并将生成的通信用户描述信息存储在用户信息库中。

用户信息生成模块 120 从外界获取通信记录信息，查询特定用户标注库 103 并判断其中是否存储有通信记录中的通信对象即联系方的标识字段值，若未存储，按照用户信息传递策略中的非特定用户的信息传递策略，将满足条件的联系方标识信息添加到通信主体的通信用户描述信息中并存储在用

户信息库 121 中；若存储有，按照用户信息传递策略中的特定用户的信息传递策略，将满足条件的联系方标识信息，以及该联系方标识信息对应的存储在特定用户标注库 103 中的用户标注信息添加到通信主体的通信用户描述信息中并存储在用户信息库 121 中。

如果通信对象为特定用户标注库中的特定用户，将用户标注信息添加到用户信息库 121 中即在通信主体的通信用户描述信息中记录偏好属性，如果该偏好属性已存在，则对偏好属性信息进行调整，如增加没有的可传递参数，或更改参数对应的权值等；如果该偏好属性不存在，则增加该偏好属性信息即可。

这里，通信主体指通信记录的归属用户，比如在一条主叫话单记录中为主叫号码用户，此时通信主体由话单记录中的用户号码字段标识。

策略设置模块 123，是面向系统管理员的管理接口，用于设置各种用户信息传递策略。实际情况是不同的类具有不同特征，需要设置不同的用户信息传递策略。设置好的用户信息传递策略存储在策略池 122 中。

用户信息传递策略包括针对非特定用户的属性建立条件和特定用户的属性建立条件。这里的非特定用户是相对特定用户而言的，指在特定用户标注库 103 中不存在的用户。针对非特定用户的属性建立的条件决定是否将非特定用户加入通信主体的通信用户描述信息中，针对特定用户的属性建立的条件决定是否将特定用户加入通信主体的通信用户描述信息中。

用户信息传递策略具体内容是字段值限制条件描述的组合，这里的字段来自通信记录。一条用户信息传递策略可以用运算式描述为式 (1) 所示：

$$I = \{Q_1 \cap Q_2 \cap \dots \cap Q_n\} \quad (n \geq 1) \quad (1)$$

式 (1) 中，I 表示一条用户信息传递策略内容，Q 表示不同的字段值即限制条件， \cap 表示前后的限制条件间为和的关系，如式 (1) 所示，只用全部满足限制条件 $Q_1 \sim Q_n$ ，才认为满足条件，符合用户信息传递策略。需要说明的是，式 (1) 仅仅是用户信息传递策略的一种表达方式，前后的限制条件间也可以为或的关系，也可以部分为和的关系部分为或的关系等等，这

里不再一一列举。

表 3 给出了应用模板名称为汽车销售类和汽车维修类的特定用户的用户信息传递策略内容和对应的含义：

应用模板名称	用户信息传递策略内容	含义
汽车销售	$I_1 = \{ (\text{通话次数} > 1) \}$	只要满足通话次数大于一次的汽车销售联系方的可传递信息，都会被记录在通信主体的通信用户描述信息中
汽车维修	$I_2 = \{ (\text{通话次数} > 3) \cap (\text{通话总时长} > 00:15:00) \}$	只有满足通话次数大于三次，并且通话总时长超过十五分钟的汽车维修联系方的可传递信息，才会被记录在通信主体的通信用户描述信息中

表 3

策略池 122，用于存储各种用户信息传递策略，并提供给用户信息生成模块使用。在策略池 122 中，一个应用模板名称对应一条用户信息传递策略，用户信息生成模块 120 根据原始资源数据所属的应用模板名称查询并调用策略池 120 中对应的用户信息传递策略。

用户信息库 121，用于存储通信主体的用户描述信息文件，包括但不限于用户的标识 ID、用户的社会关系属性、偏好属性。

社会关系属性中记录满足社会关系建立条件的非特定用户和特定用户及与其的联系权值，联系权值是一个表示用户之间关系度的参数，可以根据特定字段经过一定的组合运算得出，代表通信主体与该非特定用户或特定用户的联系频繁度或亲密度。这里，以其中一种组合运算为例，假设联系权值用 S 表示，特定字段由系统设定，用 P 表示，则有式 (2)：

$$S = f(P_1, P_2, \dots, P_n) = \frac{P_m}{\sum_{i=1}^n P_i}, \quad n \geq m \geq 1 \quad (2)$$

偏好属性主要由不同应用模板的可传递参数组成，描述了已建立社会关系的特定用户的特征，这些特征在满足建立社会关系条件之后，传递给通信主体，用于描述通信主体的通信用户描述信息。表 4 是一个用户信息记录的例子，其中用户 1 表示通信主体，在该通信主体的通信用户描述信息中包括

有非特定用户 1、非特定用户 2、特定用户 1 和特定用户 2 的信息，偏好属性由应用模板 1 和应用模板 2 的可传递参数组成，描述了已建立社会关系的特定用户 1 和特定用户 2 的特征。

```

<用户 1>
  <号码></号码>
  <社会关系属性>
    <非特定用户>
      <非特定用户 1>
        <号码></号码>
        <联系权值></联系权值>
      </非特定用户 1>
      <非特定用户 2>
        <号码></号码>
        <联系权值></联系权值>
      </非特定用户 2>
    </非特定用户>
    <特定用户>
      <特定用户 1>
        <号码></号码>
        <名称></名称>
        <联系权值></联系权值>
      </特定用户 1>
      <特定用户 2>
        <号码></号码>
        <名称></名称>
        <联系权值></联系权值>
      </特定用户 2>
    </特定用户>
  </社会关系属性>
  <偏好属性>
    <应用模板 1>
      <模板名称>汽车销售</模板名称>
      <可传递参数 a></可传递参数 a>
      <可传递参数 b></可传递参数 b>
    </应用模板 1>
    <应用模板 2>
      <模板名称>游艇销售</模板名称>
      <可传递参数 a></可传递参数 a>
      <可传递参数 b></可传递参数 b>
    </应用模板 2>
  </偏好属性>
</用户 1>

```

表 4

进一步地，如果用户信息生成模块 120 从外界获取的通信记录信息中，不存在用户信息传递策略中的限制条件中描述的字段，则本发明系统还可以包括预处理单元 11，用于对通信记录进行预处理，从原始通信记录字段中提取预先设定的目标字段组成新记录并存储，目标字段至少包括标识通信记录归属者的通信主体标识字段、标识通信对象的联系方标识字段，同时还包括按照预设要求对通信记录进行相应处理并由此产生的附加字段。

预处理单元 11 主要包括记录预处理模块 110 和新记录库 111。其中，

记录预处理模块 110，用于对从外界获取的通信记录信息进行预处理，从原始通信记录字段中提取预先设定的目标字段组成新记录并存储，同时该新记录还包括按照预设要求对通信记录进行相应处理并由此产生的附加字段。由目标字段和附加字段组成新的通信记录，存入新记录库 111 中。

这里，以对从外部获得的原始话费清单进行预处理为例进行说明，表 5 是一种原始话费清单记录，包括用户号码、序号、呼叫发生地、主被叫类型、呼叫类型、对方号码、通话日期、通话时间、通话时长、移动话费和长途话费。假设预先设定的目标字段为用户号码、序号、对方号码、通话次数和通话总时长，从表 5 可见，用户号码、序号和对方号码可以直接从原始话费清单记录中提取，而通话次数和通话总时长式原始话费清单记录中没有的，可以通过对同类话单记录进行合并来获得，即合并联系方标识字段值相同的话单记录，这里指对方号码相同的话单记录。这里的合并也就是相加，比如要获得对方号码为 A 的通话总时长，只需将原始话费清单记录中对方号码为 A 的个条记录中的通话时长相加即可。

原始话费清单记录										
用户 号码	序号	呼叫 发生地	主被叫 类型	呼叫 类型	对方 号码	通话 日期	通话 时间	通话 时长	移动 话费	长途 话费

表 5

新记录库 111，用于存储经过预处理的通信记录，用户信息生成模块 120 从新记录库 111 中获取所需的通信记录。记录字段由记录预处理模块 110 为

预设的目标字段。按照上例中，经过预处理存入新记录库 111 的话单记录如表 6 所示：

预处理后的通信记录					
用户 号码	序号	呼叫 发生地	主被叫 类型	通话 次数	通话 总时长

表 6

表 6 中，用户号码字段表示该话单记录是该用户号码对应的通信用户的话单，通话次数指的是同类话单条目总数，通话总时长指同类话单的通话时长之和。

以上详细描述了本发明生成通信用户描述信息的系统，下面结合图 1，具体描述本发明生成通信用户描述信息的方法，图 2 是本发明提供的实施例中生成通信用户描述信息的方法的流程图，该方法可以由预设定时器控制执行，或者按照后台管理员下发命令执行，主要包括以下步骤：

步骤 200：预先设置用户信息传递策略。

用户信息传递策略包括针对非特定用户的属性建立条件和特定用户的属性建立条件，针对非特定用户的属性建立的条件决定是否将非特定用户相关信息加入通信主体的通信用户描述信息中，针对特定用户的属性建立的条件决定是否将特定用户相关信息加入通信主体的通信用户描述信息中。关于信息传递策略，在描述策略设置模块 123 时已详细介绍，这里不再重述。

步骤 201：获取通信记录信息，判断通信记录信息中的联系方标识信息是否存在于生成的用户标注信息中，若存在，进入步骤 203；否则进入步骤 202。

获取通信记录的方法可以通过很多手段获得，如从一个专门存储通信记录的数据库中读取等，具体实现与本发明无关，这里不再详述。

步骤 202：按照预设的用户信息传递策略中的非特定用户的信息传递策略，将满足条件的通信记录信息中的联系方标识信息记录到通信主体的通信用户描述信息中，之后结束本流程。

联系方标识可以是电话号码等。

步骤 203: 按照预设的用户信息传递策略中的特定用户的信息传递策略, 将满足条件的通信记录信息中的联系方标识信息, 以及该联系方标识信息对应的用户标注信息记录到通信主体的通信用户描述信息中。

本步骤中, 联系方标识信息对应的用户标注信息, 可以是在步骤 201 中判断出通信记录信息中的联系方标识信息存在于生成的用户标注信息中时, 从生成的用户标注信息中获得的可传递信息, 也可以是在步骤 203 中判断出满足条件后, 依据该联系方标识从生成的用户标注信息中获得的可传递信息。

所述记录的用户标注信息的方法为: 遍历所述通信主体用户的用户描述信息中预设的偏好属性, 判断该偏好属性中是否存在所述用户标注信息所属的类, 若不存在, 则在该通信主体用户的偏好属性中添加所述用户标注信息所属的类及其中的可传递参数; 若存在, 则进一步判断该偏好属性中是否存在所述用户标注信息中的可传递参数, 若存在, 结束; 若不存在, 则将所述用户标注信息中的可传递参数添加在该偏好属性中。

本步骤中, 用户标注信息的生成方法如图 3 所示, 图 3 是本发明提供的实施例中实现信息标注的流程图, 包括以下步骤:

步骤 300: 提取一条原始资源数据记录, 获取该记录的所属类的类名称。这里, 一条原始资源数据记录指通信记录信息中的一条记录。

步骤 301: 根据获得的类名称, 调用对应该类名称的应用模板, 并根据调用的应用模板中包括的字段, 将该记录中对应的字段值记录在该应用模板中。

本步骤中, 应用模板是预先设置好的。

当出现应用模板中的部分字段在原始资源数据记录中没有对应字段时, 可以有下面两种处理方式:

方式一, 忽略该字段, 将应用模板中该字段的值设置为空, 该字段的值可以由企业或管理员通过自服务平台进行手工配置。

方式二，应用模板中的某些字段在原始资源数据记录中没有明确的对应字段，但是原始资源数据记录中的某些字段内容可能包含有应用模板中该字段需要的信息，这里，将原始资源数据记录中可能包含有应用模板中该字段需要的信息的某些字段称为模糊字段。图4是本发明提供的实施例中一种记录用户应用模板的方法的流程图，包括以下步骤：

步骤400：提取应用模板中的字段。

步骤401：原始资源数据记录中是否存在该字段，若存在，则进入步骤405；否则，进入步骤402。

步骤402：原始资源数据记录中是否存在该字段的模糊字段，若存在，则进入步骤403；否则，进入步骤404。

步骤403：从模糊字段中分析提取关键字段，并将该关键字段的值记录在应用模板中，之后结束。

本步骤中可以采用现有的分词法对模糊字段进行切词分析，提取有效关键词。切词分析方法很多，具体实现属于本领域技术人员公知技术，这里不再赘述。

在这种方式下，需要预先在标注模块100中设置应用模板中某些字段与其对应的模糊字段的映射表，这样，可以通过该映射表查询到从应用模板中提取的字段对应的模糊字段。同时，在这种方式下，标注模块100还进一步包括切词分析功能。

步骤404：将应用模板中该字段的值置为空，之后结束。

步骤405：在应用模板中记录该字段的值。

下面结合表7所示的实例，给出映射表的一个示例：

应用模板中的字段	模糊字段
品牌	公司简介
口味	公司简介

表 7

假设原始资源数据记录中的一条数据记录了一家汽车销售公司的注册

信息，该注册信息包括该公司的电话号码、公司名称、所属行业、地址、公司简介，公司简介中有“宝马”字样描述。调用的汽车销售应用模板上的参数有电话号码、公司名称、地址、品牌，其中品牌字段有备选值：马自达、宝马、尼桑、大众等。在进行用户信息标注时，从原始记录中提取电话号码、公司名称、地址三个字段的值写入汽车销售应用模板中的对应字段，由于品牌字段在原始资源数据记录中没有对应字段，从表7所示的映射表中得出其模糊字段为公司简介，从原始资源数据记录的公司简介字段中进行切词分析，将“宝马”定为品牌字段的值写入汽车销售应用模板中的品牌字段。

步骤 302: 存储完成记录的应用模板。

自此，实现了对用户信息的标注，是该用户称为特定用户。

步骤 303: 判断是否还有原始资源数据库，若有，则返回步骤 300；否则结束。

结合图1中对信息标注单元10的描述，可知经过标注的特定用户信息存储在特定用户标注库103中。

从图3所示的信息标注流程来看，一个很重要的过程就是调用应用模板来对特定用户进行描述。应用模板基于本体语言设计，将原始资源数据写入应用模板之后，标注模块100会自动生成该原始资源数据所描述的特定用户信息的标注信息文件，完成特定用户的属性信息标注。下面对比对两个汽车销售公司的信息进行信息标注前后的情况，表8是标注前的汽车销售公司的信息：

企业名称	电话号码	地区	地址	所属类	业务介绍
深业丰田汽车销售服务有限公司	075-26710686	广东省 深圳市	麻雀岭工业区9栋	汽车销售	主要经营 <u>丰田</u> 品牌的汽车销售，同时提供 <u>汽车护理</u> 服务
飞扬汽车销售服务有限公司	075-82391435	广东省 深圳市	麻雀岭工业区38栋	汽车销售	主要经营 <u>奇瑞</u> 等品牌汽车销售，此外还提供 <u>汽车保险</u> 服务

表 8

表8中，加有下划线的黑体字为从原始资源数据记录中的模糊字段 - 业务介绍中切出的关键词汇。

经过标注，将原始资源数据信息写入应用模板形成表9所示的标注信息：

<pre> <特定用户 1> <电话号码>0755-26710686</电话号码> <模板名称>汽车销售</模板名称> <属性参数> <组 1: 基本属性> <企业名称>深业丰田汽车销售服务有限公司</企业名称> <地区 >广东省深圳市</地区 > <地址 >麻雀岭工业区 9 栋</地址> <邮编 ></邮编 > <网址 ></网址 > </组 1: 基本属性> <组 2: 扩展属性 > <品牌 >丰田</品牌 > <服务项 >汽车销售、汽车护理</服务项 > </组 2: 扩展属性 > </属性参数> </特定用户 1> <特定用户 2> <电话号码>0755-82391435</电话号码> <模板名称>汽车销售</模板名称> <属性参数> <组 1: 基本属性> <企业名称>飞扬汽车销售服务有限公司</企业名称> <地区 >广东省深圳市</地区> <地址>麻雀岭工业区 38 栋</地址> <邮编 ></邮编 > <网址 ></网址 > </组 1: 基本属性> <组 2: 扩展属性 > <品牌 >奇瑞</品牌 > <服务项 >汽车销售、汽车保险</服务项 > </组 2: 扩展属性 > </属性参数> </特定用户 2> </pre>

表 9

表9中，斜体字为标注出的信息。

进一步地，如果获取的通信记录信息中不存在用户信息传递策略中的限制条件中描述的字段，则本发明方法还可以包括对通信记录进行预处理，即

从原始通信记录字段中提取预先设定的目标字段组成新记录并存储,目标字段至少包括标识通信记录归属者的通信主体标识字段、标识通信对象的联系方标识字段,同时新记录还包括按照预设要求对通信记录进行相应处理并由此产生的附加字段。

目标字段的设定依据通信记录的格式不同而不同,由管理员设定。下面结合一个对原始话费清单进行预处理为例进行说明预处理的方法。

假设某用户 13526843612 六月前十天的话费清单如表 10 所示:

序号	用户号码	呼叫发生地	主被叫类型	呼叫类型	对方号码	通话日期	通话时间	通话时长	移动电话费	长途话费
1	13526843612	755	主叫	本地	13343857303	2006-6-1	17: 13	00: 12: 23	/	/
2	13526843612	755	主叫	本地	13343857303	2006-6-1	23: 00	00: 45: 56	/	/
3	13526843612	755	被叫	本地	13343857303	2006-6-1	19: 00	00: 09: 48	/	/
4	13526843612	755	主叫	本地	0755-26710686	2006-6-5	09: 30	00: 15: 33	/	/
5	13526843612	755	主叫	本地	0755-26710686	2006-6-8	14: 30	00: 05: 53	/	/
6	13526843612	755	主叫	本地	0755-82391435	2006-6-8	15: 30	00: 03: 53	/	/
7	13526843612	755	主叫	本地	0755-82391435	2006-6-9	08: 30	00: 05: 53	/	/
8	13526843612	755	被叫	本地	13892983773	2006-6-9	10: 30	00: 02: 26	/	/
9	13526843612	755	被叫	本地	13343857303	2006-6-10	12: 00	00: 09: 12	/	/

表 10

假设预处理针对原始话费清单的格式,设定的目标字段包括用户号码、序号、对方号码即联系方标识字段,所述按照要求对通信记录进行相应处理为合并对方号码相同的话单记录,增加的附加字段包括通话次数、通话总时长。经过合并处理后形成的新记录如表 11 所示:

用户号码	序号	对方号码	通话次数	通话总时长
13526843612	1	13343857303	4	01: 17: 19
13526843612	2	13892983773	1	00: 02: 26
13526843612	3	0755-26710686	2	00: 21: 26

13526843612	4	0755-82391435	2	00: 09: 46
-------------	---	---------------	---	------------

表 11

假设可以以某种手段获取通话距离的话,也可以将通话距离作为一个参数加以考虑,用于判断通信双方的关系度。

以下结合图 1 举例说明用户信息生成过程,图 5 是实施例的流程图,包括以下步骤:

步骤 500: 判断是否存在用户通信记录,若存在,进入步骤 501; 否则结束。

这里假设,用户信息生成模块 120 从新记录库 111 中提取了一组预处理话单,如表 12 所示,其中用户号码为通信主体标识字段,对方号码为联系方标识字段。因此,本步骤中判断的结果是存在用户通信记录。

用户号码	序号	对方号码	通话次数	通话总时长
13526843612	1	13343857303	4	01: 17: 19
13526843612	2	13892983773	1	00: 02: 26
13526843612	3	0755-26710686	2	00: 21: 26
13526843612	4	0755-82391435	2	00: 09: 46

表 12

步骤 501 ~ 步骤 503: 获取一条用户通信记录,并从该记录中获取通信主体标识,并判断是否存在该通信主体标识对应的通信主体的信息,若存在,则进入步骤 504; 若不存在,则创建一条该通信主体的用户信息记录。

本步骤中,假设用户信息生成模块 120 提取的一条用户通信记录如表 13 所示。

用户号码	序号	对方号码	通话次数	通话总时长
13526843612	1	13343857303	4	01: 17: 19

表 13

用户信息生成模块 120 遍历用户信息库 121, 判断在用户信息库 121 中, 是否记录有用户号码 13526843612, 若未记录, 则创建该用户的用户信息记录文件 (如表 4), 即在用户信息库生成该用户 13526843612 的一条空的用

描述信息记录。

步骤504: 从该用户通信记录中获取联系方标识。

用户信息生成模块120读取表13所示通信记录的联系方标识号码, 即对方号码为13343857303。

步骤505: 判断是否存在该联系方的标注信息, 若存在, 则进入步骤507; 否则进入步骤506。

本步骤中, 假设用户信息生成模块120遍历特定用户标注库103, 查询出对方号码13343857303未经标注。

步骤506: 调用非特定用户的信息传递策略, 并进入步骤508。

本步骤假设用户信息生成模块120从策略池122调用的非特定用户的信息传递策略为: $I_p = \{ (\text{通话次数} > 3) \cap (\text{通话总时长} > 00:30:00) \}$ 。

步骤507: 调用特定用户的信息传递策略。

步骤508: 判断是否满足条件, 若满足, 进入步骤509; 否则, 返回步骤500。

本步骤中, 用户信息生成模块120根据该非特定用户的信息传递策略 I_p 中描述的字段, 依次读取表13所示的用户通信记录中的通话次数和通话总时长, 判断是否满足 I_p 中设定的条件。

很明显, 本实施例中, 表13中显示的通话次数为四次, 通话总时长为一小时十七分十九秒, 均满足 I_p 中设定的条件即通话次数大于三次且通话总时长大于三十分钟。

步骤509: 记录联系方的信息, 并返回步骤500。

将联系方13343857303添加至13526843612的用户描述信息中。

进一步地, 可以生成联系方的联系权值并记录。假设联系权值为一定时期内与该联系方的通话次数占总通话次数的比例, 如表12所示, 在这4条记录的时段内, 联系方13343857303的联系权值 $S = \frac{4}{4+1+2+2} = \frac{4}{9}$ 。

以上流程完成了非特定用户13343857303的信息作为通信主体

13526843621的用户描述信息的生成过程。

按照上述方法，提取第二条记录，假设查询出第二条记录中对方号码13892983773为非特定用户，采用图5所示流程可以得出，同样调用非特定用户的信息传递策略，发现该第二条记录中通话次数为一次，通话总时长为两分二十六秒，不满足Ip中设定的条件，所以忽略该话单记录，即丢弃。

提取第三条记录，假设查询出第三条记录中对方号码0755-26710686为特定用户，联系方0755-26710686所属类为汽车销售类，该联系方的标注信息如表9所示。采用图5所示流程可以得出，在步骤507中，特定用户标注库103从标注信息中读取0755-26710686所属的应用模板名称，从策略池122中调用该应用模板相应的特定用户的信息传递策略。假设特定用户的信息传递策略为： $I = \{ (\text{通话次数} > 1) \}$ 。从第三条记录可见，用户号码13526843612与对方号码0755-26710686的通话次数为两次，满足I设定的条件即通话次数大于一次，因此，将联系方0755-26710686的信息添加至通信主体13526843612的用户描述信息中。这里需要说明的是，除了将对方号码0755-26710686添加至通信主体13526843612的用户描述信息中外，还同时添加对方号码0755-26710686对应的用户标注信息中的可传递参数。

进一步地，生成联系方的联系权值并记录。假设联系权值为一定时期内与该联系方的通话次数占总通话次数的比例，如表12所示，在这4条记录的时段内，联系方0755-26710686的联系权值 $S = \frac{2}{4+1+2+2} = \frac{2}{9}$ 。

由于联系方0755-26710686是经过标注的特定用户，所以记录对方号码0755-26710686对应的用户标注信息中的可传递参数的方法也就是在用户描述信息中添加相应的偏好属性。具体方法为：首先遍历该通信主体用户的偏好属性，判断该通信主体用户的偏好属性中是否有汽车销售类的属性记录，本实施例中，假设通信主体用户第一次联系汽车销售类的特定用户，所以直接在用户偏好属性中添加汽车销售类记录和相应的可传递参数，该实施例中假设汽车销售模板的可传递参数为品牌和服务项。

提取第四条记录，假设联系方 0755-82391435 所属类为汽车销售类，该联系方的标注信息如表 9 所示，处理流程同第三条记录，只是在最后偏好属性的添加过程中，由于原来已经有汽车销售类的偏好属性记录，所以处理这条记录时，需要判断该偏好属性记录中是否存在联系方 0755-82391435 的标注信息中的可传递参数，若存在，则不做任何操作，或者将联系方 0755-82391435 的标注信息中的可传递参数添加并覆盖原来的记录；若不存在，则将联系方 0755-82391435 的标注信息中的可传递参数添加在该偏好属性中，以更新偏好属性。

根据图5所示的流程，对表12中的四条记录处理完成后，最终可以生成用户的描述信息如表14所示：

```

<用户 1>
  <号码>13526843612</号码>
  <社会关系属性>
    <非特定用户>
      <非特定用户 1>
        <号码>13343857303</号码>
        <联系权值>4/9</联系权值>
      </非特定用户 1>
    </非特定用户>
    <特定用户>
      <特定用户 1>
        <号码>0755-26710686</号码>
        <名称>深业丰田汽车销售服务有限公司</名称>
        <联系权值>2/9</联系权值>
      </特定用户 1>
      <特定用户 2>
        <号码>0755-82391435</号码>
        <名称>飞扬汽车销售服务有限公司</名称>
        <联系权值>2/9</联系权值>
      </特定用户 2>
    </特定用户>
  </社会关系属性>
  <偏好属性>
    <应用模板 1>
      <模板名称>汽车销售</模板名称>
      <品牌 >丰田、奇瑞</品牌 >
      <服务项>汽车销售、汽车护理、汽车保险</服务项>
    </应用模板 1>
  </偏好属性>

```

</用户 1>

表 14

以上仅以话费清单作为实施例进行描述，本发明方案对于Email、IM等通信方式同样适用，如将企业Email与IM号码进行标注。另外，本发明方案对于彩铃下载等数据业务的处理也适用，比如可以为各类彩铃制作相应的本体描述模板，将服务提供商（SP）提供的各类彩铃作为特定用户进行标注，然后从用户数据业务的服务记录中分析用户描述信息。

以上所述，仅为本发明的较佳实施例而已，并非用于限定本发明的保护范围，凡在本发明的精神和原则之内所做的任何修改、等同替换、改进等，均应包含在本发明的保护范围之内。

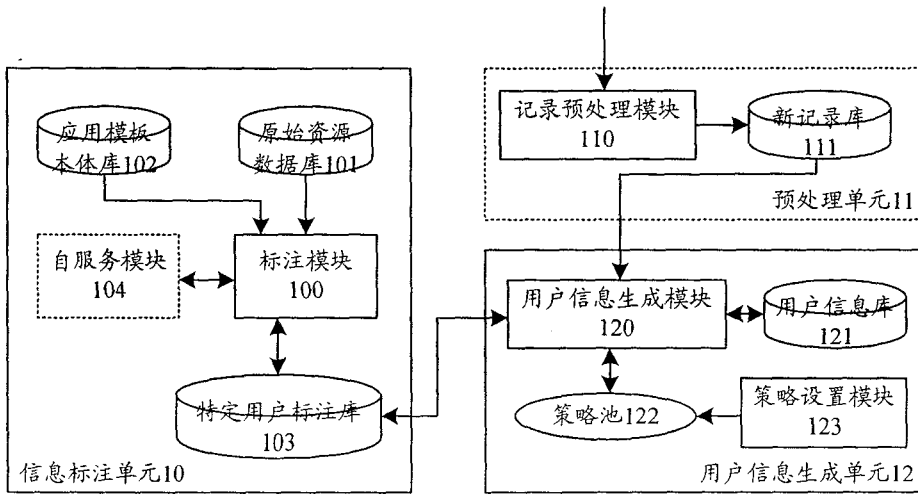


图 1

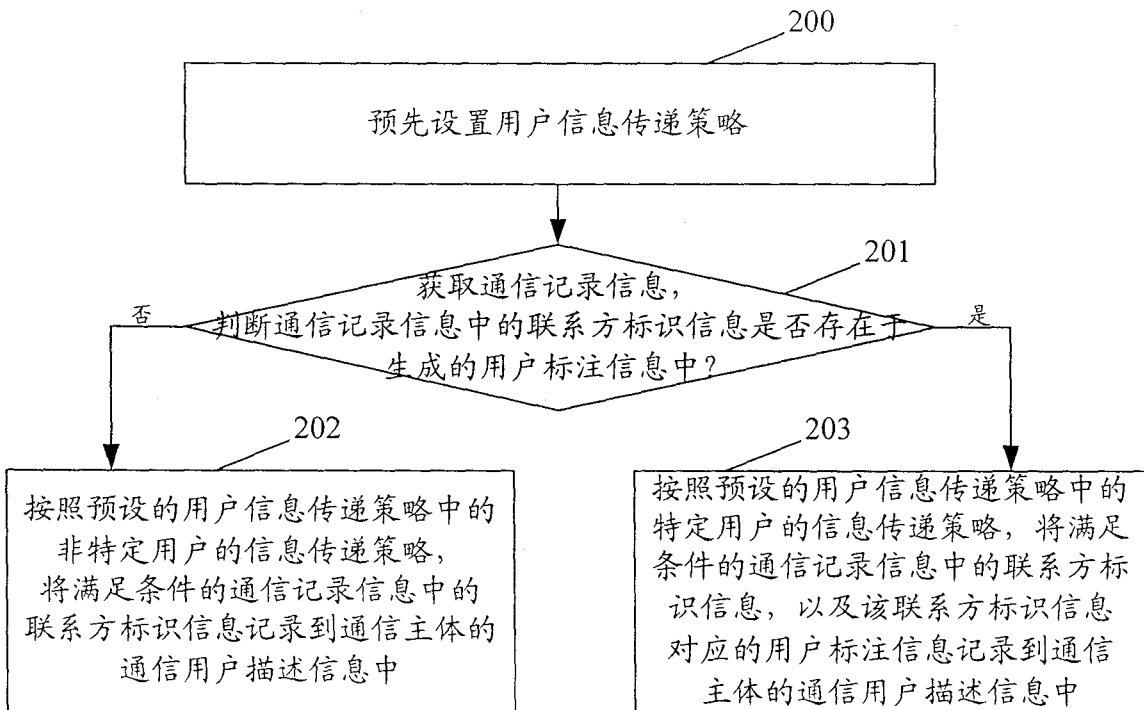


图 2

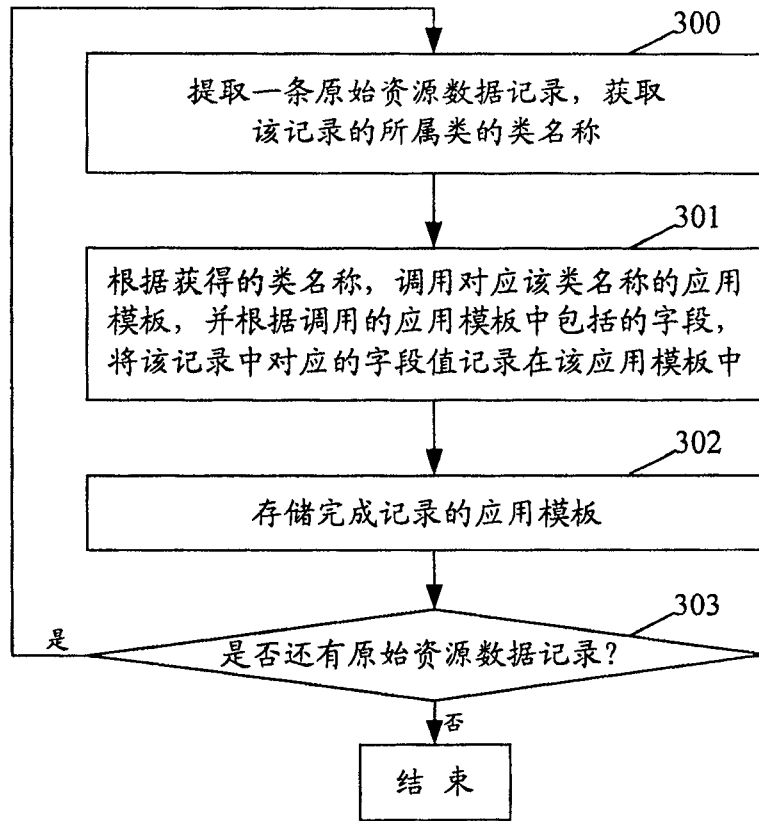


图 3

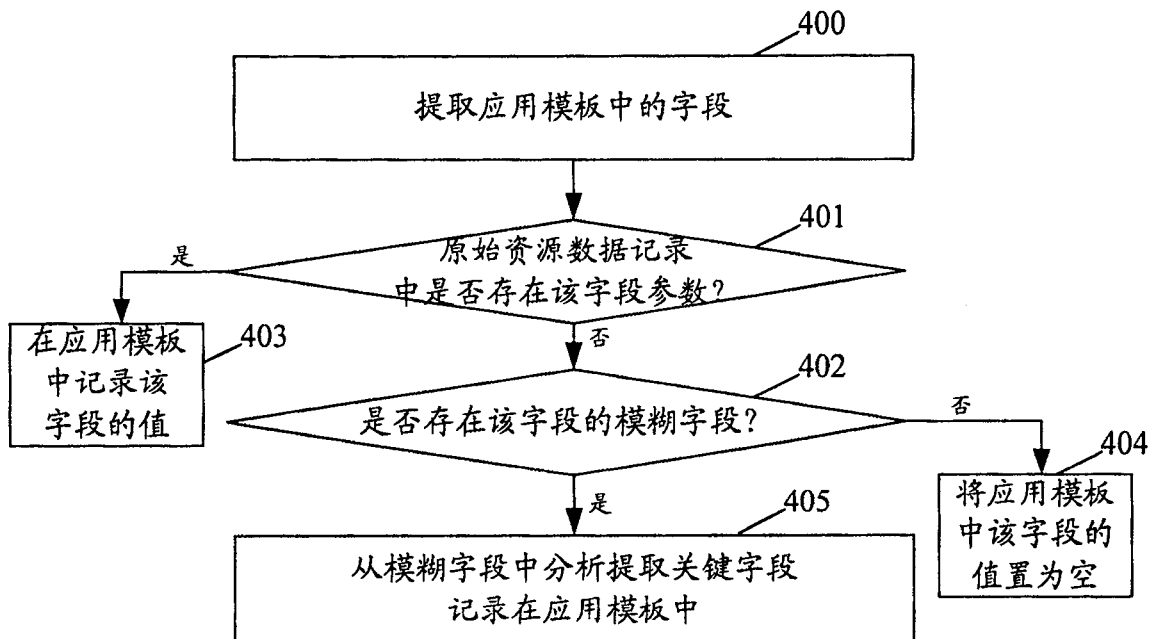


图 4

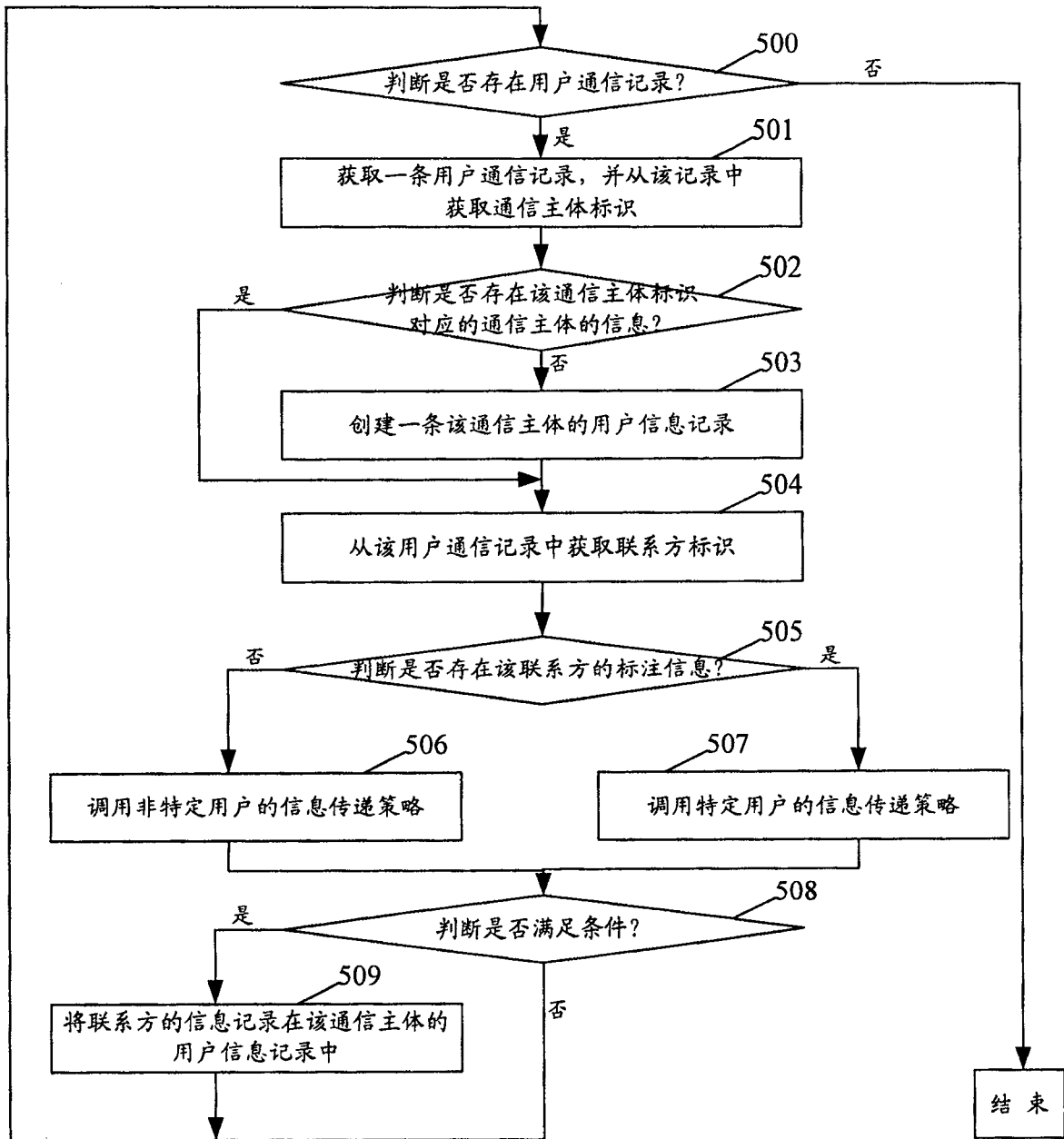


图 5