

# 一种针对不平衡微博用户数据中异常用户检测方法

张子牛，王军杰，李涤非

指导老师：刘震

2022 年 8 月 31 日

本发明涉及信息技术领域，尤其涉及一种针对不平衡微博用户数据集中个体异常用户和集体异常用户的检测；首先通过模式挖掘方法 (*Pattern Mining Algorithm*) 学习数据的全局模式；遍历每一个样本点，根据学习到全局模式对该点的标签，针对数据不平衡的特点对其进行损失函数值 (*Loss Function Value*) 的修正；将修正后的结果融入 *GNN* 编码器的特征学习过程中进行异常检测。

说明书摘要附图如图 1 所示：

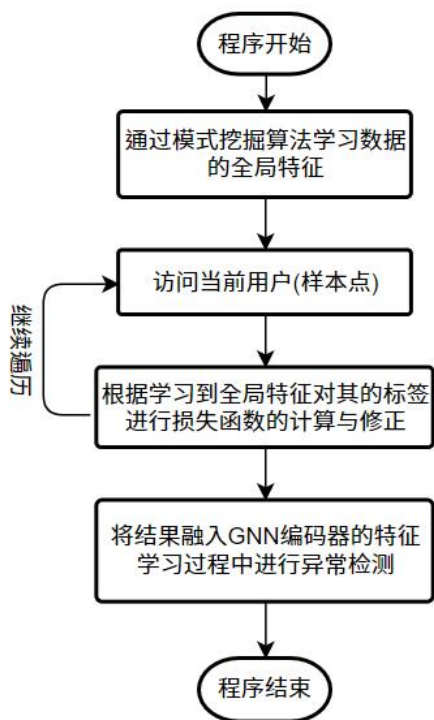


图 1: 说明书摘要流程图

## 一种不平衡微博用户数据中集体异常用户检测方法

1. 一种不平衡微博用户数据中集体异常用户检测方法，其特征在于，包括以下步骤：

- 步骤 1：学习已有的不平衡微博数据集的全局模式，根据建成的图找到图中的密度块，并据此对用户节点打上标签；
- 步骤 2：针对每一个用户节点选出特定数目与其连接的其他用户节点；
- 步骤 3：针对每一个用户节点，根据整体数据集非平衡的性质，对损失函数进行修正；
- 步骤 4：利用图编码器聚合各个用户节点的特征信息对其分类，识别出异常用户节点。

2. 根据权利要求 1 所述的一种不平衡微博用户数据中集体异常用户检测方法，其特征在于，所述步骤 1 包括以下内容：

- 根据原本具有不平衡性质的微博用户数据集建立二分图，图的一侧为用户，另一侧为属性；
- 在二分图上使用图模式挖掘算法 (*Data Mining Algorithm*)，先识别出二分图中密度较大的聚集块，然后针对用户节点与不同聚集块的相对关系，学习得到一个  $n$  行的矩阵， $n$  为用户节点个数，矩阵的第  $i$  行表示第  $i$  个用户节点的独热向量编码，根据独热向量编码的特征对每一个用户打上标签： $\{0, 1\}$ 。

3. 根据权利要求 1 所述的一种不平衡微博用户数据中集体异常用户检测方法，其特征在于，所述步骤 2 包括以下内容：

- 我们定义一个用户节点的二跳节点为该用户节点在图中向外走两步到达的节点。结合二分图的性质，一个用户节点的二跳节点为另一个用户节点。
- 考虑到整张图的容量特别大，其中用户节点的数目为 8403。对于每一个用户节点，如果我们遍历其所有的二跳邻居，则其时间复杂度为  $O(N^2)$ ，这对于大型图来说是不可接受的。因此我们选择以数学期望的形式通过随机游走算法 (*Random Walk Algorithm*) 随机选出每个用户节点周围 100 个二跳邻居，其中与该节点标签相同的 50 个，标签不同的 50 个。如果一个用户节点不存在 100 个二跳邻居，则将其所有的二跳邻居都选出。

4. 根据权利要求 1 所述的一种不平衡微博用户数据中集体异常用户检测方法，其特征在于，步骤 3 包括以下内容：

- 由于整体数据集并不平衡, 对于一个多数类用户节点, 它周围的点绝大多数都是与之标签相同的, 反之对于一个少数类用户节点, 它周围的点绝大多数都是与之标签不同的。二者在选择相邻节点的时候都是随机选择的, 因此对于少数类用户节点, 为了提高判断的准确度, 我们需要适当增大与相邻节点标签类型判断错误的惩罚。同时对于多数类用户节点, 我们需要适当减小与相邻节点标签类型判断错误的惩罚, 因为本身来说与当前用户节点标签相同的点有很多。可以发现, 这个惩罚的大小是与当前用户节点标签类型的数量相关, 因此, 我们提出了集体异常用户检测时不平衡数据修正系数  $\alpha_{coe1}, \alpha_{coe2}$ 。

- 我们给出如下定义:

定义当前遍历到的用户节点:  $u$ ;

定义与当前用户节点  $u$  标签相同的用户节点:  $u_{pos}$ ;

定义用户集合:  $U$ , 即  $u, u_{pos}, u_{neg} \in U$ ;

定义少数类标签用户节点的数目:  $n_n$ ;

定义多数类标签用户节点的数目:  $n_x$ ;

定义当前用户节点  $u$  的标签:  $lab_u$ ;

定义标签为  $lab_u$  的点的数目:  $n_{lab_u}$ ;

定义用户节点  $u$  与用户节点  $v$  的相似程度:  $simi(u, v)$ ;

定义用户节点  $u$  的低维表示:  $\Phi_u$ 。

- 传统意义上, 基于随机游走遍历到节点的损失函数多用铰链损失函数 (*Hinge Loss Function*) 来表示:

$$Loss_{HingeLoss}(u) = \mathbb{E}_{u_{pos} \in U_{pos}, u_{neg} \in U_{neg}} \max\{0, \Delta_{Bias} - \Phi_u \times \Phi_{u_{pos}}^T + \Phi_u \times \Phi_{u_{neg}}^T\}$$

- 针对此微博数据中不平衡的特点, 以及集体异常用户检测与个体异常用户检测的区别, 我们对传统铰链损失函数进行修正, 提出了集体异常用户检测时不平衡数据修正系数  $\alpha_{coe1}, \alpha_{coe2}$  以及如下损失函数公式:

$$Loss(u) = \mathbb{E}_{u_{pos} \in U_{pos}, u_{neg} \in U_{neg}} \max\{0, \Theta - \alpha_{coe1} \cdot simi(u, u_{pos}) - \alpha_{coe2} \cdot simi(u, u_{neg})\}$$

其中,  $\Theta = \frac{\vartheta}{n_{lab_u}^{1/4}}$ ,  $\vartheta$  为一个给定的超参数。

如果当前用户点的标签是少数类标签, 则  $\alpha_{coe1} = 1 + \frac{n_n}{n_n + n_x}$ ,  $\alpha_{coe2} = 1 + \frac{n_n}{n_n + n_x}$ , 如果当前用户点的标签是多数类标签, 则  $\alpha_{coe1} = 1 - \frac{n_n}{n_n + n_x}$ ,  $\alpha_{coe2} = 1 - \frac{n_n}{n_n + n_x}$ 。

5. 根据权利要求 1 所述的一种不平衡微博用户数据中集体异常用户检测方法, 其特征在于, 步骤 4 包括以下内容:

- 该步骤使用基于图卷积神经网络 (*Graph Neural Network*) 的编码器聚合了每一个用户节点的特征信息, 以此对用户节点进行分类, 判断识别出的聚集块是否异常, 同时识别出集体异常用户。

## 一种不平衡微博用户数据中个体异常用户检测方法

1. 一种不平衡微博用户数据中集体异常用户检测方法，其特征在于，包括以下步骤：

- 步骤 1：学习已有的不平衡微博数据集的全局模式，并对用户节点打上标签；
- 步骤 2：针对每一个用户节点选出特定数目与其连接的其他用户节点；
- 步骤 3：针对每一个用户节点，根据整体数据集非平衡的性质，对损失函数进行修正；
- 步骤 4：利用图编码器聚合各个用户节点的特征信息对其分类，识别出异常用户节点。

2. 根据权利要求 1 所述的一种不平衡微博用户数据中个体异常用户检测方法，其特征在于，所述步骤 1 包括以下内容：

- 根据原本具有不平衡性质的微博用户数据集建立二分图，图的一侧为用户，另一侧为属性；
- 在二分图上使用图模式挖掘算法 (*Data Mining Algorithm*)，针对用户节点的个体特征，学习得到一个  $n$  行的矩阵， $n$  为用户节点的个数，矩阵的每  $i$  行表示用户节点  $i$  的独热向量编码，根据独热向量编码的特征对每一个用户打上标签： $\{0, 1\}$ 。

3. 根据权利要求 1 所述的一种不平衡微博用户数据中个体异常用户检测方法，其特征在于，所述步骤 2 包括以下内容：

- 我们定义一个用户节点的二跳节点为该用户节点在图中向外走两步到达的节点。结合二分图的性质，一个用户节点的二跳节点为另一个用户节点。
- 考虑到整张图的容量特别大，其中用户节点的数目为 8403。对于每一个用户节点，如果我们遍历其所有的二跳邻居，则其时间复杂度为  $O(N^2)$ ，这对于大型图来说是不可接受的。因此我们选择以数学期望的形式通过随机游走算法 (*Random Walk Algorithm*) 随机选出每个用户节点周围 100 个二跳邻居，其中与该节点标签相同的 50 个，标签不同的 50 个。如果一个用户节点不存在 100 个二跳邻居，则将其所有的二跳邻居都选出。

4. 根据权利要求 1 所述的一种不平衡微博用户数据中个体异常用户检测方法，其特征在于，步骤 3 包括以下内容：

- 由于整体数据集并不平衡，对于一个多数类用户节点，它周围的点绝大多数都是与之标签相同的，反之对于一个少数类用户节点，它周围的点绝大多数都是与之标签不同的。二者在选择相邻节点的时候都是随机选择的，因此对于少数类用户节点，为了提高判断的准确度，我们需要适当增大与相邻

节点标签类型判断错误的惩罚。同时对于多数类用户节点，我们需要适当减小与相邻节点标签类型判断错误的惩罚，因为本身来说与当前用户节点标签相同的点有很多。可以发现，这个惩罚的大小是与当前用户节点标签类型的数量相关。在实验过程中我们发现，由于个体异常检测步骤 1 的过程中使用具体的图挖掘算法不同，有的图挖掘算法在打完标签后生成的数据并没有过于不平衡（例如少数类数目：多数类数目  $\approx 1:3$ ）。综上，我们提出了个体异常用户检测时不平衡数据修正系数  $\gamma_{coe1}, \gamma_{coe2}$  及偏置系数  $\beta$ 。

- 我们给出如下定义：

定义当前遍历到的用户节点： $u$ ；

定义与当前用户节点  $u$  标签相同的用户节点： $u_{pos}$ ；

定义用户集合： $U$ ，即  $u, u_{pos}, u_{neg} \in U$ ；

定义少数类标签用户节点的数目： $n_n$ ；

定义多数类标签用户节点的数目： $n_x$ ；

定义当前用户节点  $u$  的标签： $lab_u$ ；

定义标签为  $lab_u$  的点的数目： $n_{lab_u}$ ；

定义用户节点  $u$  与用户节点  $v$  的相似程度： $simi(u, v)$ ；

定义用户节点  $u$  的低维表示： $\Phi_u$ 。

- 传统意义上，基于随机游走遍历到节点的损失函数多用铰链损失函数 (*Hinge Loss Function*) 来表示：

$$Loss_{HingeLoss}(u) = \mathbb{E}_{u_{pos} \in U_{pos}, u_{neg} \in U_{neg}} \max\{0, \Delta_{Bias} - \Phi_u \times \Phi_{u_{pos}}^T + \Phi_u \times \Phi_{u_{neg}}^T\}$$

- 针对此微博数据中不平衡的特点，以及个体异常用户检测与集体异常用户检测的区别，我们对传统铰链损失函数进行修正，提出了个体异常用户检测时不平衡数据修正系数  $\gamma_{coe1}, \gamma_{coe2}$  以及如下损失函数公式：

$$Loss(u) = \mathbb{E}_{u_{pos} \in U_{pos}, u_{neg} \in U_{neg}} \max\{0, \Theta - \lambda - \gamma_{coe1} \cdot simi(u, u_{pos}) - \gamma_{coe2} \cdot simi(u, u_{neg})\}$$

其中， $\Theta = \frac{\vartheta}{n_{lab_u}^{1/4}}$ ， $\vartheta, \lambda$  为两个给定的超参数， $\gamma_{coe1} = 4 \cdot (\alpha_{coe1} - \beta) + 1$ ， $\gamma_{coe2} = 4 \cdot (\alpha_{coe2} - \beta) + 1$ ，

$\beta$  为一个偏置系数，与标签的不平衡程度反相关，也即  $\beta \propto \frac{n_n + n_x}{n_n}$ 。

如果当前用户点的标签是少数类标签，则  $\alpha_{coe1} = 1 + \frac{n_n}{n_n + n_x}$ ， $\alpha_{coe2} = 1 + \frac{n_n}{n_n + n_x}$ ，如果当前用户点的标签是多数类标签，则  $\alpha_{coe1} = 1 - \frac{n_n}{n_n + n_x}$ ， $\alpha_{coe2} = 1 - \frac{n_n}{n_n + n_x}$ 。

5. 根据权利要求 1 所述的一种不平衡微博用户数据中集体异常用户检测方法，其特征在于，步骤 4 包括以下内容：

- 该步骤使用基于图卷积神经网络 (*Graph Neural Network*) 的编码器聚合了每一个用户节点的特征信息，以此对用户节点进行分类，同时识别出个体异常用户。



# 一种不平衡微博用户数据中异常用户检测方法

## 技术领域

本发明涉及信息技术领域，尤其涉及一种不平衡微博用户数据中异常用户检测方法

## 背景技术

在当前最普及最具有代表性的社交媒体平台微博上，网络水军等异常用户及群体的存在，往往能够对当前时事政治把控社会舆论的走向，引领民意。如果不加以识别和控制，对于社会根基和国家政治基础会有极大危害，甚至颠覆国家政权。

当前，针对微博数据集这种大规模数据集中的异常检测主要有两类，一类是根据用户节点之间结构和拓扑关系的图挖掘算法 (*Data Mining Algorithm*) 学习全局属性进行异常检测，另一类是使用图卷积神经网络模型 (*Graph Neural Network*) 根据节点的邻域特征进行聚合，从而实现异常节点检测。考虑到图异常并不具有图挖掘中基于随机游走算法 (*Random Walk Algorithm*) 的全局属性，因此现有图卷积神经网络模型在图异常检测的基准测试中表现欠佳。同时由于忽略了节点个体本身的特征，认为所有点的属性具有相似性，同样对图异常损失造成了偏差。有学者已经提出了将模式挖掘与特征学习结合起来，将图挖掘算法学习到的全局模式融入到图卷积神经网络编码器的特征学习过程中的方法，对问题做出了改进。

然而，由于微博用户数据具有比较显著的不平衡的特性，异常用户节点通常只占整个用户节点很小的一部分，同时对于一个少类用户节点群体边缘的节点，其邻居节点中绝大多数与自身并不属于同一类别，因此对其进行分类的时候不能对每个用户节点一视同仁进行计算，有学者提出一种通过加偏置的方法进行修正，不过加偏置并不能从根本上对同类型与不同类型用户节点之间的相似度进行调整。数据不平衡问题对整体分类效果产生的负面影响依旧存在。

## 发明内容

为了解决上述现有技术中存在的问题，本发明提供了一种不平衡微博用户数据中异常用户检测方法，拟解决目前由于微博数据不平衡而对整体分类效果产生负面影响的问题。其中包括个体异常用户检测方法和集体异常用户检测方法。

一种不平衡微博用户数据中异常集体用户检测方法，包括以下步骤：

步骤 1：学习已有的不平衡微博数据集的全局模式，根据建成的图找到图中的密度块，并据此对用户节点打上标签；

步骤 2：针对每一个用户节点，选出特定数目与其连接的其他用户节点；

步骤 3：针对每一个用户节点，根据整体数据集非平衡的性质，对损失函数进行修正；

步骤 4：利用图编码器，聚合各个用户节点的特征信息并对其分类，识别出异常用户节点。

优选的，所述步骤 1 包括以下内容：

根据原本具有不平衡性质的微博用户数据集建立二分图，图的一侧为用户，另一侧为属性标签；

在二分图上使用图模式挖掘算法 (*Data Mining Algorithm*), 先识别出二分图中密度较大的聚集块, 然后针对用户节点与不同聚集块的相对关系, 学习得到一个  $n$  行的矩阵,  $n$  为用户节点个数, 矩阵的每一行为一个独热向量, 根据独热向量的特征对每一个用户打上标签:  $\{0, 1\}$ .

优选的, 所述步骤 2 包括以下内容:

我们定义一个用户节点的二跳节点为该用户节点在图中向外走两步到达的节点。结合二分图的性质, 一个用户节点的二跳节点为另一个用户节点。

考虑到整张图的容量特别大, 其中用户节点的数目为 8403。对于每一个用户节点, 如果我们遍历其所有的二跳邻居, 则其时间复杂度为  $O(N^2)$ , 这对于大型图来说是不可接受的。因此我们选择以数学期望的形式通过随机游走算法 (*Random Walk Algorithm*) 随机选出每个用户节点周围 100 个二跳邻居, 其中与该节点标签相同的 50 个, 标签不同的 50 个。如果一个用户节点不存在 100 个二跳邻居, 则将其所有的二跳邻居都选出。

优选的, 所述步骤 3 包括以下内容:

我们首先给出如下定义:

定义当前遍历到的用户节点:  $u$ ;

定义与当前用户节点  $u$  标签相同的用户节点:  $u_{pos}$ ;

定义用户集合:  $U$ , 即  $u, u_{pos}, u_{neg} \in U$ ;

定义少数类标签用户节点的数目:  $n_n$ ;

定义多数类标签用户节点的数目:  $n_x$ ;

定义当前用户节点  $u$  的标签:  $lab_u$ ;

定义标签为  $lab_u$  的点的数目:  $n_{lab_u}$ ;

定义用户节点  $u$  与用户节点  $v$  的相似程度:  $simi(u, v)$ ;

定义用户节点  $u$  的低维表示:  $\Phi_u$ 。

传统意义上, 基于随机游走遍历到节点的损失函数多用铰链损失函数 (*Hinge Loss Function*) 来表示:

$$Loss_{HingeLoss}(u) = \mathbb{E}_{u_{pos} \in U_{pos}, u_{neg} \in U_{neg}} \max\{0, \Delta_{Bias} - \Phi_u \times \Phi_{u_{pos}}^T + \Phi_u \times \Phi_{u_{neg}}^T\}$$

针对此微博数据中不平衡的特点, 以及集体异常用户检测与个体异常用户检测的区别, 我们对传统铰链损失函数进行修正, 提出了集体异常用户检测时不平衡数据修正系数  $\alpha_{coe1}, \alpha_{coe2}$  以及如下损失函数公式:

$$Loss(u) = \mathbb{E}_{u_{pos} \in U_{pos}, u_{neg} \in U_{neg}} \max\{0, \Theta - \alpha_{coe1} \cdot simi(u, u_{pos}) - \alpha_{coe2} \cdot simi(u, u_{neg})\}$$

其中,  $\Theta = \frac{\vartheta}{n_{lab_u}^{1/4}}$ ,  $\vartheta$  为一个给定的超参数。

如果当前用户点的标签是少数类标签, 则  $\alpha_{coe1} = 1 + \frac{n_n}{n_n + n_x}$ ,  $\alpha_{coe2} = 1 + \frac{n_n}{n_n + n_x}$ , 如果当前用户点的标签是多数类标签, 则  $\alpha_{coe1} = 1 - \frac{n_n}{n_n + n_x}$ ,  $\alpha_{coe2} = 1 - \frac{n_n}{n_n + n_x}$ 。

优选的, 步骤 4 包括以下内容:

该步骤使用基于图卷积神经网络 (*Graph Neural Network*) 的编码器聚合了每一个用户节点的特征信息, 以此对用户节点进行分类, 判断识别出的聚集块是否异常, 同时识别出集体异常用户。

一种不平衡微博用户数据中个体异常用户检测方法, 包括以下步骤:

步骤 1: 学习已有的不平衡微博数据集的全局模式, 并对用户节点打上标签;

步骤 2: 针对每一个用户节点选出特定数目与其连接的其他用户节点;

步骤 3: 针对每一个用户节点, 根据整体数据集非平衡的性质, 对损失函数进行修正;

步骤 4: 利用图编码器聚合各个用户节点的特征信息对其分类, 识别出异常用户节点。

优选的, 步骤 1 包括以下内容:

根据原本具有不平衡性质的微博用户数据集建立二分图, 图的一侧为用户, 另一侧为属性标签;

在二分图上使用图模式挖掘算法 (*Data Mining Algorithm*), 针对用户节点的个体特征, 学习得到一个  $n$  行的矩阵,  $n$  为用户节点的个数, 矩阵的每  $i$  行表示用户节点  $i$  的独热向量编码, 根据独热向量编码的特征对每一个用户打上标签:  $\{0, 1\}$ .

优选的, 步骤 2 包括以下内容:

我们定义一个用户节点的二跳节点为该用户节点在图中向外走两步到达的节点。结合二分图的性质, 一个用户节点的二跳节点为另一个用户节点。

考虑到整张图的容量特别大, 其中用户节点的数目为 8403。对于每一个用户节点, 如果我们遍历其所有的二跳邻居, 则其时间复杂度为  $O(N^2)$ , 这对于大型图来说是不可接受的。因此我们选择以数学期望的形式通过随机游走算法 (*Random Walk Algorithm*) 随机选出每个用户节点周围 100 个二跳邻居, 其中与该节点标签相同的 50 个, 标签不同的 50 个。如果一个用户节点不存在 100 个二跳邻居, 则将其所有的二跳邻居都选出。

优选的, 步骤 3 包括以下内容:

我们首先给出如下定义:

定义当前遍历到的用户节点:  $u$ ;

定义与当前用户节点  $u$  标签相同的用户节点:  $u_{pos}$ ;

定义用户集合:  $U$ , 即  $u, u_{pos}, u_{neg} \in U$ ;

定义少数类标签用户节点的数目:  $n_n$ ;

定义多数类标签用户节点的数目:  $n_x$ ;

定义当前用户节点  $u$  的标签:  $lab_u$ ;

定义标签为  $lab_u$  的点的数目:  $n_{lab_u}$ ;

定义用户节点  $u$  与用户节点  $v$  的相似程度:  $simi(u, v)$ ;

定义用户节点  $u$  的低维表示:  $\Phi_u$ 。

传统意义上, 基于随机游走遍历到节点的损失函数多用铰链损失函数 (*Hinge Loss Function*) 来表示:

$$Loss_{HingeLoss}(u) = \mathbb{E}_{u_{pos} \in U_{pos}, u_{neg} \in U_{neg}} \max\{0, \Delta_{Bias} - \Phi_u \times \Phi_{u_{pos}}^T + \Phi_u \times \Phi_{u_{neg}}^T\}$$

针对此微博数据中不平衡的特点, 以及个体异常用户检测与集体异常用户检测的区别, 我们对传统铰链损失函数进行修正, 提出了个体异常用户检测时不平衡数据修正系数  $\gamma_{coe1}, \gamma_{coe2}$  以及如下损失函数公式:

$$Loss(u) = \mathbb{E}_{u_{pos} \in U_{pos}, u_{neg} \in U_{neg}} \max\{0, \Theta - \lambda - \gamma_{coe1} \cdot simi(u, u_{pos}) - \gamma_{coe2} \cdot simi(u, u_{neg})\}$$

其中,  $\Theta = \frac{\vartheta}{n_{lab_u}^{1/4}}$ ,  $\vartheta, \lambda$  为两个给定的超参数,  $\gamma_{coe1} = 4 \cdot (\alpha_{coe1} - \beta) + 1$ ,  $\gamma_{coe2} = 4 \cdot (\alpha_{coe2} - \beta) + 1$ ,  $\beta$  为一个偏置系数, 与标签的不平衡程度反相关, 也即  $\beta \propto \frac{n_n + n_x}{n_n}$ 。

如果当前用户点的标签是少数类标签, 则  $\alpha_{coe1} = 1 + \frac{n_n}{n_n + n_x}$ ,  $\alpha_{coe2} = 1 + \frac{n_n}{n_n + n_x}$ , 如果当前用户点的标签是多数类标签, 则  $\alpha_{coe1} = 1 - \frac{n_n}{n_n + n_x}$ ,  $\alpha_{coe2} = 1 - \frac{n_n}{n_n + n_x}$ 。

优选的, 步骤 4 包括以下内容:

该步骤使用基于图卷积神经网络 (Graph Neural Network) 的编码器聚合了每一个用户节点的特征信息, 以此对用户节点进行分类, 同时识别出个体异常用户。

本发明的有益效果包括:

1. 对于集体异常检测, 本发明针对不平衡微博数据集, 提出了损失函数计算过程中的修正系数  $\alpha_{coe1}$  和  $\alpha_{coe2}$ , 有效解决了不平衡数据对集体异常检测结果造成的负面影响。
2. 对于个体异常检测, 本发明针对不平衡微博数据集, 提出了损失函数计算过程中的修正系数  $\gamma_{coe1}$ ,  $\gamma_{coe2}$  以及偏置系数  $\beta$ , 有效解决了不平衡数据对个体异常检测结果造成的负面影响。
3. 本发明识别精度高, 方便监管者对于微博或其他社交媒体平台海量数据集的异常检测监管。

## 附图说明

图 2 为本发明提出的一种针对不平衡微博用户数据中异常用户检测方法的流程图。

图 3 为集体异常用户检测中由图挖掘算法得到的独热向量编码。

图 4 为修正系数的图形化表示。

图 5 为已有方法与本发明提供的方法结果对比。

## 具体实施方式

为使本申请实施例的目的、技术方案和优点更加清楚, 下面将结合本申请实施例中附图, 对本申请实施例中的技术方案进行清楚、完整地描述, 显然, 所描述的实施例仅是本申请的一部分实施例, 而不是全部的实施例。通常在此处附图中描述和示出的本申请实施例的组件可以各种不同的配置来布置和设计。因此, 以下对在附图中提供的本申请的实施例的详细描述并非旨在限制要求保护本申请的范围, 而是仅仅表示本申请的选定实施例。基于本申请的实施例, 本领域技术人员在没有做出创造性劳动的前提下所获得的所有其他实施例, 都属于本申请的保护范围。

下面结合附图 1 到附图 3 对本发明的具体实施例做详细的说明:

一种不平衡微博用户数据中集体异常用户检测方法, 包括以下步骤:

步骤 1: 学习已有的不平衡微博数据集的全局模式, 根据建成的图找到图中的密度块, 并据此对用户节点打上标签;

所述步骤 1 包括以下内容:

本次使用的微博数据集是基于腾讯微博平台的一个“用户-标签”二部图数据集。该数据集共有 8405 个用户和 61964 个属性标签。该二部图中每条边上的权重表示该用户发布特定属性标签的次数。该数据集对于每个用户还提供基于时态标记出的由独热编码组成的维度为 300 的特征向量, 用以表示该用户的位置, 关键词等状态信息。

在二分图上使用图模式挖掘算法 (Data Mining Algorithm) 实现对原始用户节点的低维表示, 先识别出二分图中密度较大的聚集块, 然后针对用户节点与不同聚集块的相对关系, 学习得到一个  $n$  行的矩

阵, 如图 2 所示,  $n$  为用户节点个数, 矩阵的每一行为一个独热向量, 根据独热向量的特征对每一个用户打上标签:  $\{0, 1\}$ .

步骤 2: 针对每一个用户节点, 选出特定数目与其连接的其他用户节点;

所述步骤 2 包括以下内容:

我们定义一个用户节点的二跳节点为该用户节点在图中向外走两步到达的节点。结合二分图的性质, 一个用户节点的二跳节点为另一个用户节点。

考虑到整张图的容量特别大: 用户节点的数目为 8403。对于每一个用户节点, 如果我们遍历其所有的二跳邻居, 则其时间复杂度为  $O(N^2)$ , 这对于大型图来说是不可接受的。因此我们选择以数学期望的形式通过随机游走算法 (*Random Walk Algorithm*) 随机选出每个用户节点周围 100 个二跳邻居, 其中与该节点标签相同的 50 个, 标签不同的 50 个。如果一个用户节点不存在 100 个二跳邻居, 则将其所有的二跳邻居都选出。如图 4 所示。

步骤 3: 针对每一个用户节点, 根据整体数据集非平衡的性质, 对损失函数进行修正;

所述步骤 3 包括以下内容:

如图 5 所示, 由于整体数据集并不平衡, 对于一个多数类用户节点, 它周围的点绝大多数都是与之标签相同的, 反之对于一个少数类用户节点, 它周围的点绝大多数都是与之标签不同的。二者在选择相邻节点的时候都是随机选择的, 因此对于少数类用户节点, 为了提高判断的准确度, 我们需要适当增大与相邻节点标签类型判断错误的惩罚。同时对于多数类用户节点, 我们需要适当减小与相邻节点标签类型判断错误的惩罚, 因为本身来说与当前用户节点标签相同的点有很多。可以发现, 这个惩罚的大小是与当前用户节点标签类型的数量相关, 因此, 我们提出了集体异常用户检测时不平衡数据修正系数  $\alpha_{coe1}, \alpha_{coe2}$ 。

我们给出如下定义:

定义当前遍历到的用户节点:  $u$ ;

定义与当前用户节点  $u$  标签相同的用户节点:  $u_{pos}$ ;

定义用户集合:  $U$ , 即  $u, u_{pos}, u_{neg} \in U$ ;

定义少数类标签用户节点的数目:  $n_n$ ;

定义多数类标签用户节点的数目:  $n_x$ ;

定义当前用户节点  $u$  的标签:  $lab_u$ ;

定义标签为  $lab_u$  的点的数目:  $n_{lab_u}$ ;

定义用户节点  $u$  与用户节点  $v$  的相似程度:  $simi(u, v)$ ;

定义用户节点  $u$  的低维表示:  $\Phi_u$ 。

传统意义上, 基于随机游走遍历到节点的损失函数多用铰链损失函数 (*Hinge Loss Function*) 来表示:

$$Loss_{HingeLoss}(u) = \mathbb{E}_{u_{pos} \in U_{pos}, u_{neg} \in U_{neg}} \max\{0, \Delta_{Bias} - \Phi_u \times \Phi_{u_{pos}}^T + \Phi_u \times \Phi_{u_{neg}}^T\}$$

针对此微博数据中不平衡的特点, 以及集体异常用户检测与个体异常用户检测的区别, 我们对传统铰链损失函数进行修正, 提出了集体异常用户检测时不平衡数据修正系数  $\alpha_{coe1}, \alpha_{coe2}$  以及如下损失函数公式:

$$Loss(u) = \mathbb{E}_{u_{pos} \in U_{pos}, u_{neg} \in U_{neg}} \max\{0, \Theta - \alpha_{coe1} \cdot simi(u, u_{pos}) - \alpha_{coe2} \cdot simi(u, u_{neg})\}$$

其中,  $\Theta = \frac{\vartheta}{n_{lab_u}^{1/4}}$ ,  $\vartheta$  为一个给定的超参数。

如果当前用户点的标签是少数类标签, 则  $\alpha_{coe1} = 1 + \frac{n_n}{n_n + n_x}$ ,  $\alpha_{coe2} = 1 + \frac{n_n}{n_n + n_x}$ , 如果当前用户点的标签是多数类标签, 则  $\alpha_{coe1} = 1 - \frac{n_n}{n_n + n_x}$ ,  $\alpha_{coe2} = 1 - \frac{n_n}{n_n + n_x}$ 。

步骤 4: 利用图编码器, 聚合各个用户节点的特征信息并对其分类, 识别出异常用户节点。

所述步骤 4 包括以下内容:

该步骤使用基于图卷积神经网络 (Graph Neural Network) 的编码器聚合了每一个用户节点的特征信息, 以此对用户节点进行分类, 同时识别出个体异常用户。

在集体异常用户检测的实现中, 我们分别使用了 *Fraudar* 和 *catchsync* 两种图挖掘算法, 并通过计算 *F1* 值来衡量最终的结果, 如图 7 和表 1 所示:

表 1: 集体异常用户检测结果

Data Mining Algorithm	Frame	pre	rec	F1 score
<i>Fraudar</i>	Initial	0.8973	0.8310	0.8629
	Amended	0.8985	0.8415	0.8691
<i>Catchsync</i>	Initial	0.92	0.8099	0.8614
	Amended	0.9053	0.8415	0.8723

结合图表可以看出, 针对两种数据挖掘算法, 使用了修正后损失函数的架构对准确率均有所提升。

一种不平衡微博用户数据中个体异常用户检测方法, 包括以下步骤:

步骤 1: 学习已有的不平衡微博数据集的全局模式, 并对用户节点打上标签;

所述步骤 1 包括以下内容:

根据原本具有不平衡性质的微博用户数据集建立二分图, 图的一侧为用户, 另一侧为属性标签;

在二分图上使用图模式挖掘算法 (Data Mining Algorithm), 针对用户节点的个体特征, 学习得到一个  $n$  行的矩阵,  $n$  为用户节点的个数, 矩阵的每  $i$  行表示用户节点  $i$  的独热向量编码, 如图 6 所示。然后根据独热向量编码的特征对每一个用户打上标签:  $\{0, 1\}$ 。

步骤 2: 针对每一个用户节点选出特定数目与其连接的其他用户节点;

所述步骤 2 包括以下内容:

我们定义一个用户节点的二跳节点为该用户节点在图中向外走两步到达的节点。结合二分图的性质, 一个用户节点的二跳节点为另一个用户节点。

考虑到整张图的容量特别大, 其中用户节点的数目为 8403。对于每一个用户节点, 如果我们遍历其所有的二跳邻居, 则其时间复杂度为  $O(N^2)$ , 这对于大型图来说是不可接受的。因此我们选择以数学期望的形式通过随机游走算法 (Random Walk Algorithm) 随机选出每个用户节点周围 100 个二跳邻居, 其中与该节点标签相同的 50 个, 标签不同的 50 个。如果一个用户节点不存在 100 个二跳邻居, 则将其所有的二跳邻居都选出。

步骤 3: 针对每一个用户节点, 根据整体数据集非平衡的性质, 对损失函数进行修正;

所述步骤 3 包括以下内容：

如图 5 所示，由于整体数据集并不平衡，对于一个多数类用户节点，它周围的点绝大多数都是与之标签相同的，反之对于一个少数类用户节点，它周围的点绝大多数都是与之标签不同的。二者在选择相邻节点的时候都是随机选择的，因此对于少数类用户节点，为了提高判断的准确度，我们需要适当增大与相邻节点标签类型判断错误的惩罚。同时对于多数类用户节点，我们需要适当减小与相邻节点标签类型判断错误的惩罚，因为本身来说与当前用户节点标签相同的点有很多。可以发现，这个惩罚的大小是与当前用户节点标签类型的数量相关。在实验过程中我们发现，由于个体异常检测步骤 1 的过程中使用具体的图挖掘算法不同，有的图挖掘算法在打完标签后生成的数据并没有过于不平衡（例如少数类数目：多数类数目  $\approx 1:3$ ）。综上，我们提出了个体异常用户检测时不平衡数据修正系数  $\gamma_{coe1}, \gamma_{coe2}$  及偏置系数  $\beta$ 。

我们给出如下定义：

定义当前遍历到的用户节点： $u$ ；

定义与当前用户节点  $u$  标签相同的用户节点： $u_{pos}$ ；

定义用户集合： $U$ ，即  $u, u_{pos}, u_{neg} \in U$ ；

定义少数类标签用户节点的数目： $n_n$ ；

定义多数类标签用户节点的数目： $n_x$ ；

定义当前用户节点  $u$  的标签： $lab_u$ ；

定义标签为  $lab_u$  的点的数目： $n_{lab_u}$ ；

定义用户节点  $u$  与用户节点  $v$  的相似程度： $simi(u, v)$ ；

定义用户节点  $u$  的低维表示： $\Phi_u$ 。

传统意义上，基于随机游走遍历到节点的损失函数多用铰链损失函数 (Hinge Loss Function) 来表示：

$$Loss_{HingeLoss}(u) = \mathbb{E}_{u_{pos} \in U_{pos}, u_{neg} \in U_{neg}} \max\{0, \Delta_{Bias} - \Phi_u \times \Phi_{u_{pos}}^T + \Phi_u \times \Phi_{u_{neg}}^T\}$$

针对此微博数据中不平衡的特点，以及个体异常用户检测与集体异常用户检测的区别，我们对传统铰链损失函数进行修正，提出了个体异常用户检测时不平衡数据修正系数  $\gamma_{coe1}, \gamma_{coe2}$  以及如下损失函数公式：

$$Loss(u) = \mathbb{E}_{u_{pos} \in U_{pos}, u_{neg} \in U_{neg}} \max\{0, \Theta - \lambda - \gamma_{coe1} \cdot simi(u, u_{pos}) - \gamma_{coe2} \cdot simi(u, u_{neg})\}$$

其中， $\Theta = \frac{\vartheta}{n_{lab_u}^{1/4}}$ ， $\vartheta, \lambda$  为两个给定的超参数， $\gamma_{coe1} = 4 \cdot (\alpha_{coe1} - \beta) + 1$ ， $\gamma_{coe2} = 4 \cdot (\alpha_{coe2} - \beta) + 1$ ， $\beta$  为

一个偏置系数，与标签的不平衡程度反相关，也即  $\beta \propto \frac{n_n + n_x}{n_n n_x}$ 。

如果当前用户点的标签是少数类标签，则  $\alpha_{coe1} = 1 + \frac{n_n}{n_n + n_x}$ ， $\alpha_{coe2} = 1 + \frac{n_n}{n_n + n_x}$ ，如果当前用户点的标签是多数类标签，则  $\alpha_{coe1} = 1 - \frac{n_n}{n_n + n_x}$ ， $\alpha_{coe2} = 1 - \frac{n_n}{n_n + n_x}$ 。

步骤 4：利用图编码器聚合各个用户节点的特征信息对其分类，识别出异常用户节点。

所述步骤 4 包括以下内容：

该步骤使用基于图卷积神经网络 (Graph Neural Network) 的编码器聚合了每一个用户节点的特征信息，以此对用户节点进行分类，同时识别出个体异常用户。

在个体异常用户检测的实现中，我们分别使用了 *REV2*，*Feagle* 和 *Lockinfer<sub>r</sub>* 三种图挖掘算法，并通过计算 *F1* 值来衡量最终的结果，如图 8 和表 2 所示：

表 2: 集体异常用户检测结果

Data Mining Algorithm	Frame	pre	rec	F1 score
<i>REV2</i>	Initial	0.8581	0.8732	0.8656
	Amended	0.9129	0.8486	0.8796
<i>Feagle</i>	Initial	0.8407	0.8732	0.8566
	Amended	0.9023	0.8451	0.8727
<i>Lockinfer<sub>r</sub></i>	Initial	0.8755	0.8169	0.8452
	Amended	0.8702	0.8732	0.8717

结合图表可以看出，针对三种数据挖掘算法，使用了修正后损失函数的架构对准确率均有所提升。

综上所述，本发明有效解决了微博数据集中由于不平衡数据而对个体异常用户检测和集体异常用户检测产生负面影响的问题。具有识别精度高的特点，方便监管者对于微博或其他社交媒体平台海量数据集的异常检测监管。

以上所述实施例仅表达了本申请的具体实施方式，其描述较为具体和详细，但不能因此理解为对本申请保护范围的限制。应当指出的是，对于本领域的普通技术人员来说，在不脱离本申请技术方案构思的前提下，还可以作出若干变形和改进，这些都属于本申请的保护范围。



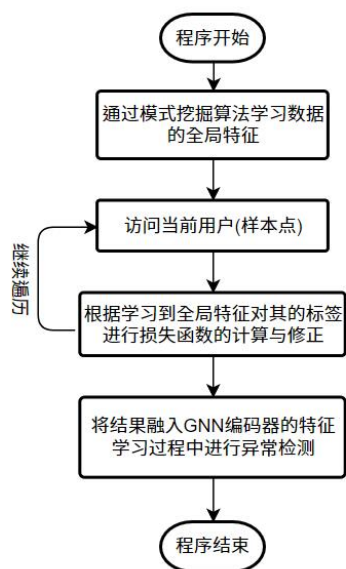


图 2: 一种针对不平衡微博用户数据中异常用户检测方法的流程图

	密度块 1	密度块 2	密度块 3	.....	密度块 m
用户节点 1:	[ 0,	1,	0,	.....	0]
用户节点 2:	[ 1,	1,	0,	.....	0]
用户节点 3:	[ 0,	1,	1,	.....	0]
用户节点 4:	[ 0,	0,	1,	.....	1]
.....					
用户节点 n:	[ 1,	0,	1,	.....	0]

图 3: 集体异常用户检测中由图挖掘算法得到的独热向量编码

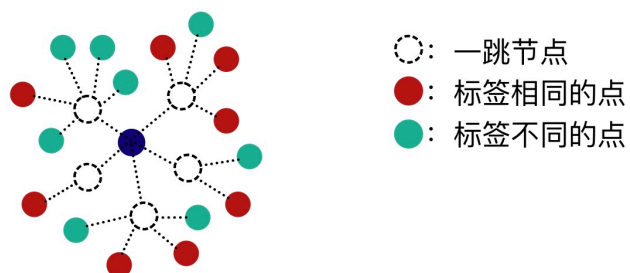


图 4: 选择二跳邻居



图 5: 少数类用户节点与多数类用户节点周围的节点分布

	特征 1	特征2
用户节点 1: [	0,	1]
用户节点 2: [	1,	0]
用户节点 3: [	1,	0]
.....		
用户节点 n: [	0,	1]

图 6: 个体异常用户检测中由图挖掘算法得到的独热向量编码

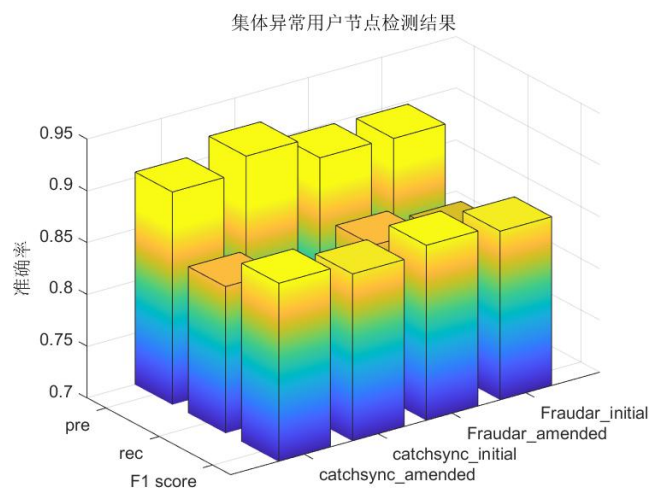


图 7: 集体异常用户节点检测结果

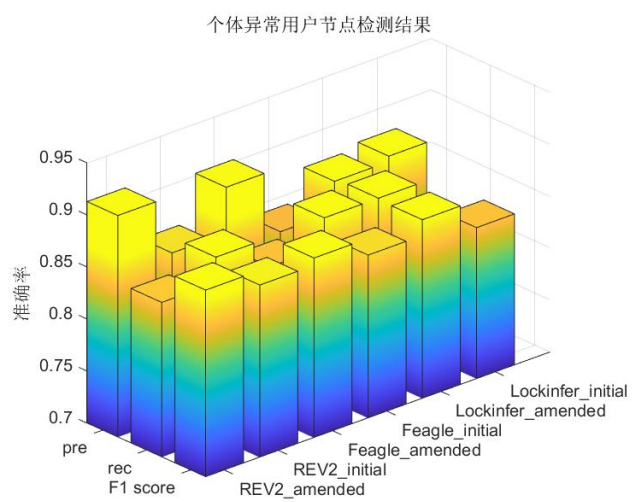


图 8: 个体异常用户节点检测结果