



(12) 发明专利

(10) 授权公告号 CN 107369081 B

(45) 授权公告日 2021.07.27

(21) 申请号 201710592343.5

审查员 乔忠华

(22) 申请日 2017.07.19

(65) 同一申请的已公布的文献号

申请公布号 CN 107369081 A

(43) 申请公布日 2017.11.21

(73) 专利权人 无锡企业征信有限公司

地址 214000 江苏省无锡市滨湖区金融一街8号17楼

(72) 发明人 宋朝钦 王安静

(74) 专利代理机构 无锡市汇诚永信专利代理事

务所(普通合伙) 32260

代理人 张欢勇

(51) Int. Cl.

G06Q 40/00 (2012.01)

权利要求书2页 说明书3页 附图1页

(54) 发明名称

用数据来源的动态影响因子确定数据有效性的系统及方法

(57) 摘要

本发明公开了一种用数据来源的动态影响因子确定数据有效性的系统:包括采集模块、预处理模块、选择模块、处理模块一、输入模块二和输出模块。方法为获取数据项多个来源的数据值;将该数据项每个数据值的来源的影响因子分别相加,得出数据项每个数据值的总影响因子;取数据项每个数据值的总影响因子最大的值为有效值;有效值的每个数据值的来源的历史有效个数+1;数据项的历史有效值个数+1;更新数据项每个来源的影响因子为新影响因子,新影响因子等于数据项该来源的历史有效个数/数据项历史有效值个数;新影响因子确定该数据项各个数据来源新的数据值到来时的有效性。保证企业信用评估数据的真实性以及全面性,提高评估结果的可信度。



1. 一种用数据来源的动态影响因子确定数据有效性的方法,其特征在於:包括:
Step1: 获取数据项多个来源的数据值;
Step2: 将该数据项每个数据值的来源的影响因子分别相加,得出该数据项每个数据值的总影响因子;
Step3: 取该数据项每个数据值的总影响因子最大的值为有效值;
Step4: 该有效值的每个数据值的来源的历史有效个数+1;该数据项的历史有效值个数+1;
Step5: 更新该数据项每个来源的影响因子为新影响因子;新影响因子等于该数据项该来源的历史有效个数/该数据项历史有效值个数;
Step6: 根据新影响因子确定该数据项各个数据来源新的数据值到来时的有效性。
2. 根据权利要求1所述的用数据来源的动态影响因子确定数据有效性的方法,其特征在於:数据项的获取包括:网络爬虫实时采集、政府各数据单位的专线实时对接、第三方数据交易平台定期批量提供、人工收集整理录入。
3. 根据权利要求1所述的用数据来源的动态影响因子确定数据有效性的方法,其特征在於:数据项包括:企业的工商注册信息;企业的法人、高管信息;企业的行政审批、处罚信息;企业的税务信息;企业的金融借贷信息;企业的知识产权信息;企业的供应商、客户信息。
4. 根据权利要求3所述的用数据来源的动态影响因子确定数据有效性的方法,其特征在於:数据项主要来源为:政府各主管部门:工商、税务、海关;银行机构:中国人民银行、中国工商银行、中国农业银行、中国建设银行、中国银行;类金融机构:担保、小贷;第三方数据交易平台:贵阳大数据交易中心;人工实地调查采集。
5. 根据权利要求1所述的用数据来源的动态影响因子确定数据有效性的方法,其特征在於:影响因子:是确定该数据项哪个版本是有效的依据;数据来源有很多,同一个数据项有多个来自于不同来源的版本,影响因子最终以确定以哪个版本为准。
6. 根据权利要求5所述的用数据来源的动态影响因子确定数据有效性的方法,其特征在於:一个数据项多个数据值,有些数据值来自于一个来源,有些数据值来自于两个或多个来源;每个数据值各自来源的影响因子相加得出该数据值的总影响因子。
7. 根据权利要求6所述的用数据来源的动态影响因子确定数据有效性的方法,其特征在於:设定各个来源的该数据项的影响因子的初始值,经过STEP4,STEP5自动调整优化,自动适应各数据来源的数据质量的动态变化。
8. 一种用数据来源的动态影响因子确定数据有效性的系统,其特征在於:包括:
数据采集模块: 获取数据项多个来源的数据值;
数据预处理模块: 将该数据项每个数据值的来源的影响因子分别相加,得出该数据项每个数据值的总影响因子;
数据选择模块: 取该数据项每个数据值的总影响因子最大的值为有效值;
数据处理模块一: 该有效值的每个数据值的来源的历史有效个数+1;该数据项的历史有效值个数+1;
数据输入模块二: 嵌套在数据处理模块一内,更新该数据项每个来源的影响因子为新影响因子;新影响因子等于该数据项该来源的历史有效个数/该数据项历史有效值个数;

数据输出模块:根据新影响因子确定该数据项各个数据来源新的数据值到来时的有效性后输出结论。

9.根据权利要求8所述的用数据来源的动态影响因子确定数据有效性的系统,其特征在于:影响因子:是确定该数据项哪个版本是有效的依据;数据来源有很多,同一个数据项有多个来自于不同来源的版本,影响因子最终以确定以哪个版本为准。

10.根据权利要求8所述的用数据来源的动态影响因子确定数据有效性的系统,其特征在于:一个数据项多个数据值,有些数据值来自于一个来源,有些数据值来自于两个或多个来源;每个数据值各自来源的影响因子相加得出该数据值的总影响因子。

用数据来源的动态影响因子确定数据有效性的系统及方法

技术领域

[0001] 本发明涉及企业征信技术领域,尤其涉及一种用数据来源的动态影响因子确定数据有效性的系统及方法。

背景技术

[0002] 随着社会的发展进步,企业信用与企业经济利益的相互关联度越来越密切。企业信用可以由征信机构通过企业信用评估手段得出,即提供信用信息服务的企业,按一定规则合法采集企业、个人的信用信息后,加工整理形成企业、个人的信用报告等征信产品。有偿提供给经济活动中的贷款方、赊销方、招标方、出租方、保险方等有合法需求的信用使用者,为其了解交易对方的信用状况提供便利。征信服务既可为防范信用风险,保障交易安全创造条件,又可使具有良好信用记录的企业和个人得以较低的交易成本获得较多的交易机会,而缺乏良好信用记录的企业或个人则相反,从而促进形成“诚信受益,失信惩戒”的社会环境。

[0003] 征信业在促进信用经济发展和社会信用体系建设中发挥着重要的基础性作用。目前大多数的征信机构,用来评估企业信用的数据没有可信性,信用信息的可信性在一定程度上可以由企业信用信息的来源决定,企业信用信息的来源有很多,例如,基于交易平台的数据,基于社交数据、基于互联网公开数据等。企业信用信息的真实性以及全面性,对企业信用的评估结果有很大的影响,但是,通过上述途径获取的企业信用信息的真实性及全面性没有保证,从而不能对企业进行全面可信的评价,如何保证企业信用评估数据的真实性以及全面性,从而提高企业信用评估结果的可信度,是本领域亟待解决的问题。

发明内容

[0004] 为解决现在技术存在的问题,本发明提供一种用数据来源的动态影响因子确定数据有效性的方法。

[0005] 为实现上述目的,本发明采用的技术方案是:一种用数据来源的动态影响因子确定数据有效性的方法,包括:

[0006] Step1:获取数据项多个来源的数据值;

[0007] Step2:将该数据项每个数据值的来源的影响因子分别相加,得出该数据项每个数据值的总影响因子;

[0008] Step3:取该数据项每个数据值的总影响因子最大的值为有效值;

[0009] Step4:该有效值的每个数据值的来源的历史有效个数+1;该数据项的历史有效值个数+1;

[0010] Step5:更新该数据项每个来源的影响因子为新影响因子,新影响因子等于该数据项该来源的历史有效个数/该数据项历史有效值个数;

[0011] Step6:根据新影响因子确定该数据项各个数据来源新的数据值到来时的有效性。

[0012] 一种用数据来源的动态影响因子确定数据有效性的系统,包括:

- [0013] 数据采集模块:获取数据项多个来源的数据值;
- [0014] 数据预处理模块:将该数据项每个数据值的来源的影响因子分别相加,得出该数据项每个数据值的总影响因子;
- [0015] 数据选择模块:取该数据项每个数据值的总影响因子最大的值为有效值;
- [0016] 数据处理模块一:该有效值的每个数据值的来源的历史有效个数+1;该数据项的历史有效值个数+1;
- [0017] 数据输入模块二:嵌套在数据处理模块一内,更新该数据项每个来源的影响因子为新影响因子,新影响因子等于该数据项该来源的历史有效个数/该数据项历史有效值个数;
- [0018] 数据输出模块:根据新影响因子确定该数据项各个数据来源新的数据值到来时的有效性后输出结论。
- [0019] 采用本发明的技术方案后,1、改变了现有技术采用的静态确定单一来源法,即对每一个企业信用信息项,通过影响因子的动态变化来指定一个可信来源,该项的所有数据均只以该来源的数据为准,其余来源的该数据项数据不在丢弃,而当成基数作为判定总影响因子最大的值为有效值的条件。2、现有技术中对每一个企业信用信息项,单纯地指定一个来源优先级列表,依次按照优先级顺序选择来源,如果高优先级来源为空,则选择次优先级来源的数据,其余来源数据丢弃。本方案改变了上述模式,将该数据项每个数据值的来源的影响因子分别相加,得出该数据项每个数据值的总影响因子;取该数据项每个数据值的总影响因子最大的值为有效值;该有效值的每个数据值的来源的历史有效个数+1;该数据项的历史有效值个数+1;所有的数据作为一个整体来分析和判断最终数据项的影响因子,来决定该数据的有效性。影响因子自动适应每个数据项各来源质量的动态变化,保障最终数据的有效性,无需人工干预。

附图说明

- [0020] 图1是利用数据来源的动态影响因子确定数据有效性的方法的流程框图。

具体实施方式

- [0021] 下面结合附图对本发明优选的技术方案做进一步的阐述:
- [0022] 如图1所示,一种用数据来源的动态影响因子确定数据有效性的方法,包括:
- [0023] Step1:获取数据项多个来源的数据值;其中数据项的获取包括:网络爬虫实时采集、政府各数据单位的专线实时对接、第三方数据交易平台定期批量提供、人工收集整理录入。数据项包括:企业的工商注册信息;企业的法人、高管信息;企业的行政审批、处罚信息;企业的税务信息;企业的金融借贷信息;企业的知识产权信息;企业的供应商、客户信息。数据项主要来源为:政府各主管部门:工商、税务、海关等;银行机构:中国人民银行、中国工商银行、中国农业银行、中国建设银行、中国银行;类金融机构:担保、小贷;第三方数据交易平台:如贵阳大数据交易中心;人工实地调查采集。
- [0024] Step2:将该数据项每个数据值的来源的影响因子分别相加,得出该数据项每个数据值的总影响因子;影响因子:是确定该数据项哪个版本是有效的依据;数据来源有很多,同一个数据项有多个来自于不同来源的版本,影响因子最终以确定以哪个版本为准。一个

数据项多个数据值,有些数据值来自于一个来源,有些数据值来自于两个或多个来源;每个数据值各自来源的影响因子相加得出该数据值的总影响因子。

[0025] Step3:取该数据项每个数据值的总影响因子最大的值为有效值;

[0026] Step4:该有效值的每个数据值的来源的历史有效个数+1;该数据项的历史有效值个数+1;具体为设定各个来源的该数据项的影响因子的初始值,经过Step3的不断自动优化调整,总影响因子最大的值即确定为最终有效值。

[0027] Step5:更新该数据项每个来源的影响因子为新影响因子,新影响因子等于该数据项该来源的历史有效个数/该数据项历史有效值个数;

[0028] Step6:根据新影响因子确定该数据项各个数据来源新的数据值到来时的有效性。

[0029] 一种用数据来源的动态影响因子确定数据有效性的系统,包括:

[0030] 数据采集模块:获取数据项多个来源的数据值;数据预处理模块:将该数据项每个数据值的来源的影响因子分别相加,得出该数据项每个数据值的总影响因子;数据选择模块:取该数据项每个数据值的总影响因子最大的值为有效值;数据处理模块一:该有效值的每个数据值的来源的历史有效个数+1;该数据项的历史有效值个数+1;数据输入模块二:嵌套在数据处理模块一内,更新该数据项每个来源的影响因子为新影响因子,新影响因子等于该数据项该来源的历史有效个数/该数据项历史有效值个数;数据输出模块:根据新影响因子确定该数据项各个数据来源新的数据值到来时的有效性后输出结论。

[0031] 其中,影响因子:是确定该数据项哪个版本是有效的依据;数据来源有很多,同一个数据项有多个来自于不同来源的版本,影响因子最终以确定以哪个版本为准。一个数据项多个数据值,有些数据值来自于一个来源,有些数据值来自于两个或多个来源;每个数据值各自来源的影响因子相加得出该数据值的总影响因子。

[0032] 以上实施方式只为说明本发明的技术构思及特点,其目的在于让熟悉此项技术的人了解本发明的内容并加以实施,并不能以此限制本发明的保护范围,凡根据本发明精神实质所做的等效变化或修饰,都应涵盖在本发明的保护范围内。

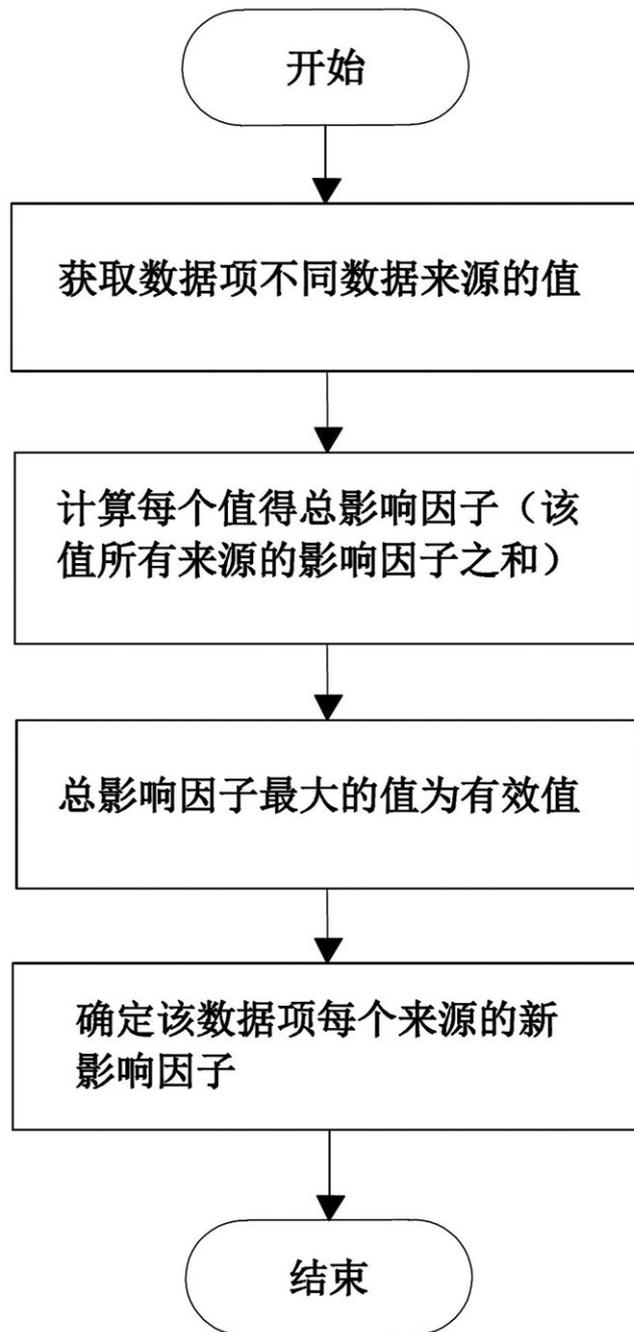


图1