可信智能计算服务

最佳实践

文档版本01发布日期2023-03-31





版权所有 © 华为技术有限公司 2024。保留一切权利。

非经本公司书面许可,任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部,并不得以任何形式传播。

商标声明

注意

您购买的产品、服务或特性等应受华为公司商业合同和条款的约束,本文档中描述的全部或部分产品、服务或 特性可能不在您的购买或使用范围之内。除非合同另有约定,华为公司对本文档内容不做任何明示或暗示的声 明或保证。

由于产品版本升级或其他原因,本文档内容会不定期进行更新。除非另有约定,本文档仅作为使用指导,本文 档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

华为技术有限公司

地址: 深圳市龙岗区坂田华为总部办公楼 邮编: 518129

网址: <u>https://www.huawei.com</u>

客户服务邮箱: <u>support@huawei.com</u>

客户服务电话: 4008302118

安全声明

漏洞处理流程

华为公司对产品漏洞管理的规定以"漏洞处理流程"为准,该流程的详细内容请参见如下网址: https://www.huawei.com/cn/psirt/vul-response-process 如企业客户须获取漏洞信息,请参见如下网址: https://securitybulletin.huawei.com/enterprise/cn/security-advisory

目录

1 基于 TICS 实现端到端的企	≥业积分查询作业1
1.1 简介	
1.2 阶段一:数据发布	
1.3 阶段二: 隐私规则防护	
1.4 阶段三:审批防护	
1.5 阶段四:基本计算能力验证	
1.6 阶段五:基于 MPC 算法的高	安全级别计算13
1.7 阶段六:统计型作业的差分隐	私保护14

基于 TICS 实现端到端的企业积分查询作业

- 1.1 简介
- 1.2 阶段一:数据发布
- 1.3 阶段二: 隐私规则防护
- 1.4 阶段三:审批防护
- 1.5 阶段四:基本计算能力验证
- 1.6 阶段五:基于MPC算法的高安全级别计算
- 1.7 阶段六:统计型作业的差分隐私保护

1.1 简介

背景信息

本案例以"小微企业信用评分"的场景为例。社保、水电气和资助金等数据统一存储 在某政务云,由不同的局进行管理,机构想单独申请进行企业相关评分的计算会非常 困难。因此可以由市政数局出面,统一制定隐私规则,审批数据提供方的数据使用申 请,并通过华为TICS可信智能计算平台进行安全计算。

图 1-1 企业信用评估应用场景示意图



数据准备

▲ 注意

- 以下数据和表结构是根据场景进行模拟的数据,并非真实数据。
- 以下数据需要提前存导入到MySQL\Hive\Oracle等用户所属数据源中,TICS本身不 会持有这些数据,这些数据会通过用户购买的计算节点进行加密计算,保障数据安 全。
- 政府信息提供方的数据tax和support,在用户计算节点agent_gov上发布。
- 能源信息提供方的数据power,在用户计算节点agent_power上发布。

表 1-1 企业税收和资助金情况表 tax

列名	含义	字段分类
Id	企业id	唯一标识
tax_bal	税收	敏感
Industry	行业类型	不敏感

表 1-2 企业政府资助金数据表 support

列名	含义	字段分类
ld	企业id	唯一标识
supp_bal	资助金的金额	敏感
Industry	行业类型	不敏感

表 1-3 企业水电情况表 power

列名	含义	字段分类
Id	企业id	唯一标识
electric_bal	电费	敏感
water_bal	水费	敏感

从业务角度考虑,安排五个阶段,来对TICS系统进行验证和测试。本章重点讲述如何 端到端实现一个该场景下的隐私计算作业完整执行流程。

3

导入数据

步骤1	在第一个合作方Partner1的MySQL数据源中,通过如下的SQL语句创建数据表:
	CREATE TABLE tax (id integer COMMENT '企业id', tax_bal integer COMMENT '税收金额', industry varchar(150) COMMENT '行业');
	CREATE TABLE support (id integer COMMENT '企业id', supp_bal integer COMMENT '资助金额', industry varchar(150) COMMENT '行业');
步骤2	在第二个合作方Partner2的MySQL数据源中,通过如下的SQL语句创建数据表: CREATE TABLE power (id integer COMMENT '企业id' , electric_bal integer COMMENT '电费', water_bal integer COMMENT '水费');
步骤3	将下面的数据分别导入csv文件并上传到MySQL数据源所在服务器。
	 Tax表的数据如下: id,tax_bal,industry 123400999,745,互联网 123400997,664,其他 123400996,243,金融 123400995,715,互联网 123400994,475,通讯 123400993,526,其他 123400992,272,互联网 123400991,646,金融 123400990,510,其他
	● Support表的数据如下:
	id,supp_bal,industry 123400999,314,互联网 123400998,405,其他 123400997,371,其他 123400996,484,金融 123400995,381,互联网 123400994,405,通讯 123400994,405,通讯 123400993,292,其他 123400991,303,互联网 123400991,303,金融 123400990,412,其他
	 Power表的数据如下: id electric bal water bal
	123400999,79,48 123400998,57,70 123400997,69,37 123400995,66,50 123400994,56,55 123400993,63,53 123400992,45,76 123400991,80,36 123400991,80,36
步骤4	执行如下SQL语句,将csv文件内的数据导入创建的数据表。
	LOAD DATA INFILE 'csv数据文件名' INTO TABLE 表名
	或者执行如下的插入语句:
	● Tax表:
文档版本 01 (2023	3-03-31) 版权所有 © 华为技术有限公司

insert into tax values (123400999,745,'互联网'), (123400998,324,'其他'), (123400997,664,'其他'), (123400996,243,'金融'), (123400995,715,'互联网'), (123400994,475,'通讯'), (123400993,526,'其他'), (123400992,272,'互联网'), (123400991,646,'金融'), (123400990,510,'其他');

• Support表:

insert into support values (123400999,314,'互联网'), (123400998,405,'其他'), (123400997,371,'其他'), (123400996,484,'金融'), (123400995,381,'互联网'), (123400994,405,'通讯'), (123400993,292,'其他'), (123400992,503,'互联网'), (123400991,303,'金融'), (123400990,412,'其他');

• Power表:

insert into power values (123400999,79,48), (123400998,57,70), (123400997,69,37), (123400996,50,57), (123400995,66,50), (123400994,56,55), (123400993,63,53), (123400992,45,76), (123400991,80,36), (123400990,39,63);

----结束

1.2 阶段一:数据发布

前提条件

完成数据准备工作。

操作步骤

步骤1 进入TICS服务控制台。

步骤2 在计算节点管理中,找到购买的计算节点,通过登录地址,进入计算节点控制台。

图 1-2 前往计算节点

5111100						
意思	基本信息					
空间管理						
空间作业	计算节点名称	agent_2959	计算节点登录地址	前往计算节点	创建时间	2024/08/31 14:19:33 GMT+08:00
通知管理	计算节点ID	4638d05696e64199a9e8b465aeea	版本类型			
计复节点管理	空间信息					
审计日志	空间区域		空间名称	test	空间ID	7
应用注册中心	部署配置					
	部署方式	云租户部署	命名空间	default	策群名称	tics-agent-4638d05696e64199a9e
	部署节点		虚拟私有云	-	子网	-
	访问IP		存储方式	OBS存储	桶名	pvc-2beffde2-9227-4eab-9fbf-25d1
	701-111,		1718/334	ODDHTM	信合	pvc-zbendez-9227-4e8b-9f0f-2501

步骤3 登录计算节点后,在下图所述位置新建连接器。

图 1-3 新建连接器

计算节点	连接器管理				
作业管理	61.82				C
任务管理	连接器名称	连接翻笑型	连接器状态	创建时间	操作
政振管理		RDS(MySql)	 正端 	2022/06/01 14:32:18 GMT+08:00	編輯 翻除
文件管理		ModelArts	 正端 	2022/05/30 10:53:34 GMT+08:00	編輯 翻除
造接器管理 (1)1日本		RDS(MySql)	 正端 	2022/05/28 20:56:11 GMT+08:00	編輯 翻除
审社管理		本地连接瞬	 正端 	2022/05/28 20:12:10 GMT+08:00	編編 翻除

步骤4 输入正确的连接信息,建立数据源和计算节点之间的安全连接。

创建连接器			
★ 连接器类型	Hive	•]
* 连接器名称	请输入]
★ Hive版本	MRS 2.x	MRS 3.x	
★ 连接地址 ⑦	多个地址请用英文分	寻分隔,	
* 用户名	请输入	<i>h</i>]
* 用户认证凭据	点击右侧按钮先添加罪	 上传	添加文件

图 1-4 输入信息

步骤5 建立完成后,连接器显示正常说明连接正常。

图 1-5 连接正常

连接器管理				
				C
连接器名称	连接關类型	连接器状态	创建时间	操作
	RDS(MySql)	● 正常	2022/06/01 14:32:18 GMT+08:00	484R 8899
	ModelArts	• 正常	2022/05/30 10:53:34 GMT+08:00	9045 (1859)
	RDS(MySqI)	• 正常	2022/05/28 20:56:11 GMT+08:00	(RE) (1950)
	本地连接翻	• 正常	2022/05/28 20:12:10 GMT+08:00	(643) 1993

步骤6进入数据管理,进行数据集发布。

图 1-6 新建数据管理

计算节点		数据管理									
作业管理	•	9/38	批量用种								С
任务管理			数编名称	连接器类型	连接器名称	数据库	数据表	状态	注册时间	操作	
数据管理				RDS趨勢	rdssadf			 已发布 	2022/06/08 12:38:03 GMT+08:00	编辑 翻译 发布	
文件管理				RDS服务	rdssadf			 已发布 	2022/06/01 09:46:27 GMT+08:00	編編 1999 没作	
连接關管理				ModelArts	macon	-	-	 已没布 	2022/05/30 14:51:57 GMT+08:00	·····································	
軍計日志				RDS服务	rdssadf			 已没布 	2022/05/28 20:56:37 GMT+08:00	明暗 1899 2015	
# 36 D AL				本地连接器	localConnector	-	-	 已没布 	2022/05/28 20:37:34 GMT+08:00	·····································	
				本地连接器	localConnector	-	-	 已没布 	2022/05/28 20:34:15 GMT+08:00	编辑 動脉 发布	

步骤7 填写参数信息。

图 1-7 墳	[写参数	ξ								
创建数据										×
基本信息										
★ 数据名称	请输入			* 连接		请选择			•	
数据描述										
				ĥ						
数据结构										
*9	字段名称	*字段类型 🎯 🔻	唯一标识 ⑦	*敏感级别 ⑦ 🙆 🔻	脱敏	2 💿 🔹	分布类型 🕐 📀 🔻	字段备注		
				6.1						



确定	取消
----	----

----结束

重复步骤1~7,发布support资助金数据表和power_data能源表。

🗋 说明

数据发布的过程并不会直接从数据源中导出用户数据,仅从数据源处获取了数据集相关的元数据 信息,用于任务的解析、验证等。

1.3 阶段二: 隐私规则防护

使用TICS的隐私规则防护能力确保数据安全。

前提条件

完成数据发布。

操作步骤

步骤1 进入多方安全计算的作业执行界面,单击创建。

图 1-8 创建作业

计算节点		多方安全计算
作业管理 ▲		A17#
多方安全计算		即建作业名称
可信联邦学习		

步骤2 在作业界面中,按照1~4提供的案例和SQL语句进行作业测试。

请输入名称 Q	C 新建 日	保存 > 执行	🖞 提交审批	ら 取消审批	▲ 格式化	⑦ SQL编写提示	
- 合作方数据							
📃 dayudayu							
+ tics222							
league_creator							
depar							
supplier_1w							
							支持文件拖拽。快
	计算过程	执行结果	审批进度	表结构			
	字段名称			类型		:	字段类别 ②
				INTEGER		I	推一标识
				STRING		:	非敏感
				STRING		:	非敏感

图 1-9 作业界面

1. 假设有人输入以下代码试图直接查询敏感数据。

select	
tax_bal, id	
from	
league_creator.tax	

×

2. 系统提示不支持进行敏感数据的SELECT操作。

图 1-10 不支持敏感操作

Privacy rule verification failed, The field type of TICS02.employee.salary is SENSITIVE, SELECT is not supported in the final result

 若试图在敏感数据中追加自己的数据,从结果倒推敏感数据,即求原数据。 Select tax_bal + electric_bal from LEAGUE_CREATOR.tax a

```
join ZZZZZZ.power_data b on a.id = b.id
```

4. TICS会识别并提示。

图 1-11 执行失败告警

C	<pre>District Werney District Content Di</pre>
	HTELITE 50796月 第88月 の 5075年20 日日日日日 日日日日日日日日日日日日日日日日日日日日日日日日日日日日日

----结束

🛄 说明

上述隐私规则,均为TICS系统提供的默认规则。

1.4 阶段三:审批防护

开启审批防护功能

前提条件

完成<mark>隐私规则防护</mark>。

操作步骤

步骤1 敏感数据被查询时,可以在审批详情中,看到是否使查询敏感数据的结果可见,可由 该提供方进行识别,并进行拒绝操作。

图 1-12 审批详情

i.

和我们在自						
92 941 m.25						
作业发起方	222222					
该计算节点执行s	ql SELECT	TAX_BAL', 'ID'	FROM LEAG	OUE_CREATOR'.'TAX'		
审批内容					-	
字段使用情况	数据集	字段名称	字段类型	是否结果中可见 ⑦	加密类型	字段作用描述
	TAX	TAX_BAL	敏感	true	園际算	
	TAX	ID	唯一标识	true	国际算	
审批意见						
审批意见						
				ر» 0/40		

步骤2 在审批详情中也可看到两个字段相加的情况,如下图所示。

图 1-13 字段相加

审批详情						
基础信息						
乍业发起方	ZZZZZZ					
亥计算节点执行so	SELECT	ID', 'TAX_BAL	FROM LEAG	BUE_CREATOR'.'TAX'	_	
审批内容						(TAX TAX BAL+(2))
字段使用情况	数据集	字段名称	字段类型	是否结果中可见 ⑦	加密类型	(1702-1702_0702-(17))
	TAX	TAX_BAL	敏感	false	国际算	(TAX.TAX_B
	TAX	ID	唯一标识	false	国际算	JOIN_ON TA
审批意见						
审批意见						
				2010		

步骤3通过查看字段是否可见,以及字段用途,能够确认该字段的应用是否符合自己的安全预期。

----结束

1.5 阶段四: 基本计算能力验证

验证TICS的基础计算能力,以计算各企业在2021年的价值评分,用于评估信贷能力, 其中的公式仅为简单的参考计算式。

前提条件

完成审批防护。

操作步骤

步骤1 执行如下的sql作业。

select c.id as `企业id`, 0.5 * a.tax_bal + 0.8 * b.supp_bal + (0.05 * c.electric_bal + 0.05 * c.water_bal) * 0.1 as `企业评分` from Partner1.TAX a, Partner1.SUPPORT b, Partner2.POWER_DATA c where b.id = c.id and a.id = b.id

步骤2 审批时可以看到如下的信息,涉及关联字段较多,其使用方式都能够在审批界面中展示出来。

图 1-14 基础信息

基础信息

作业发起方 ZZZZZZ

该计算节点执行sql SELECT 'ID', 0.5 * 'TAX_BAL' AS '10' FROM 'LEAGUE_CREATOR'. 'TAX', SELECT 'ID', 0.8 *

'SUPP_BAL' AS '10' FROM 'LEAGUE_CREATOR'.'SUPPORT'

国际算JOIN_ON TA X.TAX_BAL+0.8*SUPPORT.SUPP_ 1.05*(?)+0.05*(?))*0.1)
X.TAX_BAL+0.8*SUPPORT.SUPP_ .05*(?)+0.05*(?))*0.1)
.05*(?)+0.05*(?))*0.1)
国际算 ((0.5*TAX.TA.

步骤3执行结果如下。

图 1-15执行结果

ſ	制地通过	国保存	▶ 执行	☑ 提交审批	🗈 取消审批	卢 格式化	⑦ SQL编写提示	
1 2 3 4 5 6 7 8 9 10	select a.id 0.5 from DAYU LEAG LEAG where b.id and	as <u>fryk</u> i * b.tax_b 002.FOWER UE_CREATO UE_CREATO UE_CREATO = c.id a.id = b.	d', al + 0.8 · DATA a, R.TAX b, R.SUPPORT 1d	c.supp_ba	1 + (0.05 •	a.electric_bal	al + 0.05 * a.water_bal) * 0.1 as '企业评分'	

计算	章过程	执行结果	审批进度	表结构	
î	<u>. Nk</u> id				企业评分
12	23400585				609.535
12	23400101				645.06500000000022
12	23400647				567.665
12	23400359				557.975
12	23400821				557.639999999999994
12	23400085				578.30999999999988
12	23400927				571.7050000000022
12	23400683				560.605
12	23400565				713.004999999999977
12	23400837				572.414999999999977

步骤4 结果显示,TICS支持大量基础的SQL语法。

图 1-16 SQL 编写提醒





1.6 阶段五:基于 MPC 算法的高安全级别计算

完成demo验证阶段,为提升数据保护级别,接入以**纯密文的状态做计算**的更高安全级 别的数据,可以通过开启高隐私级别开关,提升空间安全级别。

图 1-17 高隐私级别开关

1.24.0-1.25.0		
空间详情		高级别:会默认启用高安全性的隐私计算算法 保障计算过程的安全,例如同态加密、PSI 等,但可能会影响性能以及部分作业正常执
空间名称	1.24.0-1.25.0	行。,试致别:使用国际标准的对称和4FX对称加 密结合方式,在安全沙箱内进行解密计算。性
创建人	ei_tics_	能机灾沽度致高。
描述	我随便说说 🖉	隐私保护等级 ⑦
		结果差分隐私 ⑦

再次单击作业,审批进行的同时敏感数据被进行了秘密分享加密。DAG图显示了"psi+秘密分享"的全过程流向,基本符合业界已公开的PSI算法流程和秘密分享流程。

图 1-18 加密流程



图 1-19 加密流程



1.7 阶段六:统计型作业的差分隐私保护

本示例作业,以统计各行业的"企业税收总和"与"用电量总和",进行统计分析:

Select
industry.
sum(tax hal)
sum(electric_bal)
from
LEAGUE CREATOR.tax a join
davu002 power data b
on a.id = b.id
aroup by
industry

⚠ 注意

统计分析型的作业,可能被作业执行方通过增删某个碰撞的id,得到两次作业之间的 差值,从而推算出实际taxpay和water_fee。

步骤1 开启空间中的差分隐私开关保护敏感数据,符合差分隐私条件的统计作业,会自动应 用差分隐私算法对计算结果进行加噪保护,在一定误差范围内保证数据无法被恶意偷 取。

图 1-20 差分隐私开关

0-1.25.	D	
空间详情		
空间名称	0-1.25.0	开启时,使用差分隐私算法对联邦分析作业的
创建人	ei_tics	执行结果添加隐私保护,避免历史差分攻击, 使用该功能会在发布数据集时后台计算数据统
描述	我随便说说 🖉	대통통.
		结果差分隐私 ⑦ 开)

步骤2 第一次执行作业的结果如下:

图 1-21 作业结果

〒1100000 日、保存 ▶ 执行 ♂ 提交审批 日	取消审批 占 格式化 ⑦ SQL编写提示	
<pre>1 Select 2 industry, 3 sum(electric_bal), 4 sum(electric_bal) 5 from 7 dayu002.power_data b 8 on a.id = b.id 9 group by 10 industry</pre>		
计算过程 执行结果 审批进度 表结构		
INDUSTRY	SUM(TAX_BAL)	SUM(ELECTRIC_BAL)
通讯	90061.936748504801357	10241.566257531023085
互联网	66078.857559963717677	7309.788812701260567
其他	283432.694635211473806	32114.44893234376971
金融	36593.078531096379436	4210.436015837941721

步骤3 在返回最终统计结果前,增加了一个差分隐私计算的任务节点,如图3所示。



图 1-22 差分隐私计算任务节点

```
dayu002.power_data b
on a.id = b.id
where a.id <> '123400558'
group by
industry
```

这个企业的实际tax为274:

图 1-23 tax

1 id,tax_bal,industry 2 123400558,274,互联网

得到新的结果如下:

图 1-24 新结果

联邦数据分析

作业列表 / 历史作业 / 统计各行业的税收总和和用电量总和		
	\bigcirc	
	执行成功	
结果条数 6	执行时长 14.0s	結果存於位置 /output/sqli6d9ca09270d741a8b6d9c106a359f6a7/i
INDUSTRY	SUM(TAX_BAL)	SUM(ELECTRIC_BAL)
通讯	90185.476675316059527	10258.726763214389725
互联网	66539.583321490225131	7256.443197847162061
其他	283558.757633649865894	32137.95717007242786
金融	36921.223922708768512	4218.326821712200878
餐饮	20501.52661222606329	2300.270654415101416
房地≃	26793.998662798357159	2919.710470792728171

经过计算,66539.583321490225131-66078.857559963717677=461,

通过差分隐私算法保护聚合操作的安全性,使开启算法保护的计算差值与预期得到的 实际差值274不同,避免真实数据被窃取。

-----结束