(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 119853931 A
(43)申请公布日 2025.04.18

(21)申请号 202311342115.4

(22)申请日 2023.10.17

(71)申请人 中国科学院计算机网络信息中心
地址 100190 北京市海淀区中关村南四街4号院内2号楼

(72)发明人 宫良一 付豪

(74)专利代理机构 北京知舟专利事务所(普通合伙) 11550
专利代理师 郭辗

(51)Int.Cl.
H04L 9/40(2022.01)
G06N 3/042(2023.01)
G06N 3/0442(2023.01)
G06N 3/09(2023.01)

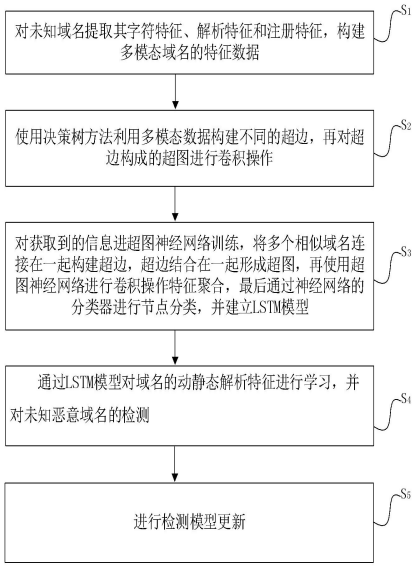
权利要求书2页 说明书5页 附图2页

(54)发明名称

基于时序超图神经网络的未知恶意域名检测方法

(57)摘要

本发明涉及域名检测技术领域,公开了基于时序超图神经网络的未知恶意域名检测方法,首先对未知域名提取其字符特征、解析特征和注册特征,构建多模态域名的特征数据;再使用决策树方法利用多模态数据构建不同的超边,再对超边构成的超图进行卷积操作,再使用超图神经网络进行卷积操作特征聚合,最后通过神经网络的分类器进行节点分类,并建立LSTM模型;再通过LSTM模型对域名的动静态解析特征进行学习,并对未知恶意域名的检测;最后进行检测模型更新。本发明结合了域名的多维特征对恶意行为进行刻画,并使用了高阶关联关系的超图模型进行检测,同时对模型进行有效更新,提高了恶意域名的检测效率和准确率。



CN 119853931 A

1.基于时序超图神经网络的未知恶意域名检测方法,其特征在于,包括以下步骤:

S_1: 对未知域名提取其字符特征、解析特征和注册特征,构建多模态域名的特征数据;

S_2:使用决策树方法利用多模态数据构建不同的超边,再对超边构成的超图进行卷积操作,对同一个域名在一段时间前后存在的多个DNS请求,多个域名一起请求,即同一个域名存在时间前后的相关性,不同域名存在相同时间访问的相关性,域名访问频率的时序特征进行卷积操作;

S_3: 对获取到的信息进超图神经网络训练,将多个相似域名连接在一起构建超边,超边结合在一起形成超图,再使用超图神经网络进行卷积操作特征聚合,最后通过神经网络的分类器进行节点分类,并建立LSTM模型;

S_4: 通过LSTM模型对域名的动静态解析特征进行学习,并对未知恶意域名的检测;

S_5:进行检测模型更新。

2.根据权利要求1所述的基于时序超图神经网络的未知恶意域名检测方法,其特征在于,在步骤S_1中,特征数据包括域名字符串长度、域名字符串的个数、域名中数字占比和域名中数字个数除以域名字符串长度域名的信息熵;和用域名中每个字符出现的频率计算出的熵值。

3.根据权利要求1所述的基于时序超图神经网络的未知恶意域名检测方法,其特征在于,其中DNS请求信息包括域名的注册人、注册域名的个体或者企业名称、域名的注册商、注册的域名商家、域名的注册时间、注册域名的日期、域名到期时间、域名到期的日期、域名的DNS服务器、用来解析域名的DNS服务器、域名的状态、域名解析的状态信息、域名的联系邮箱、域名联系人的邮箱、联系电话域名联系人的电话。

4.根据权利要求3所述的基于时序超图神经网络的未知恶意域名检测方法,其特征在于,其中对域名的注册时间、注册域名的日期、域名到期时间和域名到期的日期识别出年、月、日、时、分、秒六个特征,其他特征使用其相应的个数进行表示。

5.根据权利要求1所述的基于时序超图神经网络的未知恶意域名检测方法,其特征在于,其中DNS数据包中的域名的解析信息包括域名解析的IP信息、域名解析的别名CNAME信息、域名邮件交换服务器地址MX信息、和域名的默认TTL信息,其中域名的默认TTL信息直接使用数值,其余特征使用其相应的个数进行表示。

6.根据权利要求1所述的基于时序超图神经网络的未知恶意域名检测方法,其特征在于,在步骤S_3中,具体按以下步骤执行:

$S_{3.1}$:首先进行数据预处理,把获取到的数据分组进行编码,同一组数据当作一个整体样本进行数据处理;

$S_{3.2}$:再进行特征提取,对一组数据中提取DNS多模态特征,对提取的不同类的特征进行超边构造,分别构建决策树进行分类,把决策树分类的叶子节点的所有样本划分到相同的超边之中;

$S_{3.3}$:对不同模态特征构建的超边赋予不同权值,再从超图的矩阵中提取拉普拉斯矩阵并定义如下卷积操作,如式(1)

$$Y = Dv^{-\frac{1}{2}}HWD_e^1H^TD_v^{-\frac{1}{2}}X\Theta \quad \text{式(1)}$$

其中,X表示输入节点的特征矩阵,H为超图矩阵,D_v表示顶点度矩阵,D_e表示边的度

矩阵,Θ表示学习参数;

$S_{3.4}$:将时序行为特征与超图神经网络进行融合,设计间隔t,每隔时间t使用捕获到的数据节点构建一个超图;设置时间窗口T,T中包含多个t;判断在时间窗口T中的前后超图中是否存在相同节点,存在则对这些节点建立LSTM模型。

7.根据权利要求6所述的基于时序超图神经网络的未知恶意域名检测方法,其特征在于,在步骤$S_{3.4}$中,建立LSTM模型首先在每个时间窗口的超图进行自身的节点特征学习,再对节点进行LSTM模型的时间特征维度更新。

8.根据权利要求7所述的基于时序超图神经网络的未知恶意域名检测方法,其特征在于,建立LSTM模型时,从预处理数据中采样时间窗口为1分钟,时间总长度为1小时的数据进行训练,每个时间窗口采样5N个数据,构建大小为N的超图,在5N个数据中采用5折交叉验证方法, 其中4N数据作为训练集,剩余1N作为验证集;

模型训练完成后在线部署使用,当产生新数据时,进行实时检测恶意域名。

9.根据权利要求8所述的基于时序超图神经网络的未知恶意域名检测方法,其特征在于,其中,采取N为10000的样本进行模型训练。

10.根据权利要求1所述的基于时序超图神经网络的未知恶意域名检测方法,其特征在于,在步骤S_5中,在进行检测模型更新时,对模型的超图部分使用训练好的超图部分的权重,对模型的分类器部分进行重新训练,得到新的训练模型并进行在线部署使用。

基于时序超图神经网络的未知恶意域名检测方法

技术领域

[0001] 本发明涉及域名检测技术领域,具体涉及未知恶意域名检测方法。

背景技术

[0002] 新型互联网形态的涌现与发展给用户带来了更多的新颖网络服务,但也使得互联网安全防御变得越来越严峻。网络入侵检测是互联网安全防护的主要技术之一,通过识别网络流量内容与行为来发现网络入侵事件。由于大部分网络攻击事件往往需要利用域名来获得控制服务器地址或资源服务器地址,因此恶意域名检测成为网络入侵检测的重要内容。通过对恶意域名的检测能够尽早地发现网络入侵事件的发生。目前基于恶意域名的网络入侵检测目前常用的方案包括:(1)传统的基于统计的和规则匹配的方法,通过分析统计数据来理解知识和规则。具体而言则是利用黑白名单机制,一旦识别到恶意域名则会发出警报。(2)基于机器学习的检测方法。为了躲避黑白名单机制,恶意域名往往会采用某一些域名生成技术,例如DGA技术。通过人工提取域名的特征之后使用机器学习算法进行分类,如K近邻算法(KNN),混合聚类算法,Boosting算法等。机器学习能够帮助实现对变异类恶意域名的识别。(3)深度学习算法检测算法。然而许多高级攻击往往从多个角度来隐藏域名特征。传统的机器学习是基于结构化规则匹配,特征一旦增多往往导致欠拟合或过拟合问题。为此,深度学习算法对域名相关的所有特征进行采集,之后输入模型进行训练。此外有一些针对加密载荷直接输入网络进行训练,这类算法有多层感知机,卷积神经网络,循环神经网络,自动编码器等。

[0003] 随着新型网络的飞速发展,传统的互联网组成已由有线互联网扩展到无线互联网、物联网、车联网等。新型互联网协议与技术不仅带来了更多的网络服务,同时也给攻击者提供了更多的攻击面,导致网络防御难度大大增加。对于传统的基于恶意域名技术的网络入侵检测便是其中之一,恶意域名的生成速度与隐藏手法越来越自动化、智能化,导致传统的基于黑白名单和传统机器学习算法的恶意域名识别技术难以应对大量未知的、新型的恶意域名检测。本发现拟提出一种基于时序超图神经网络的未知恶意域名检测技术,通过对已有域名的相似性归纳分析并结合域名时序访问特征来发现潜在在互联网域名访问流量中的未知恶意域名,达到尽早发现网络入侵攻击的效果。

[0004] 机器学习和深度学习仅仅学习的是欧式空间中域名相关特征,对于已知或变种了恶意域名具有一定的识别能力,而对于未知恶意域名特征则难以有效学习。因为恶意域名特征并非严格的欧式空间结构关系而是非欧式空间的结构关系,一旦域名特征之间的关系不符合欧式空间结构关系,传统的检测方法将无法有效对其进行识别。

发明内容

[0005] 本发明的目的在于解决上述现有技术中存在的难题,提供基于时序超图神经网络的未知恶意域名检测方法。本发明结合了域名的多维特征对恶意行为进行刻画,并使用了高阶关联关系的超图模型进行检测,同时对模型进行有效更新,提高了恶意域名的检测效

率和准确率。

[0006] 本发明是通过以下技术方案实现的:包括以下步骤:

S_1: 对未知域名提取其字符特征、解析特征和注册特征,构建多模态域名的特征数据;

S_2:使用决策树方法利用多模态数据构建不同的超边,再对超边构成的超图进行卷积操作,对同一个域名在一段时间前后存在的多个DNS请求,多个域名一起请求,即同一个域名存在时间前后的相关性,不同域名存在相同时间访问的相关性,域名访问频率的时序特征进行卷积操作;

S_3: 对获取到的信息进超图神经网络训练,将多个相似域名连接在一起构建超边,超边结合在一起形成超图,再使用超图神经网络进行卷积操作特征聚合,最后通过神经网络的分类器进行节点分类,并建立LSTM模型;

S_4: 通过LSTM模型对域名的动静态解析特征进行学习,并对未知恶意域名的检测;

S_5:进行检测模型更新,在进行检测模型更新时,对模型的超图部分使用训练好的超图部分的权重,对模型的分类器部分进行重新训练,得到新的训练模型并进行在线部署使用。

[0007] 进一步的,在步骤S_1中,特征数据包括域名字符串长度、域名字符串的个数、域名中数字占比和域名中数字个数除以域名字符串长度域名的信息熵;和用域名中每个字符出现的频率计算出的熵值。

[0008] 进一步的,其中DNS请求信息包括域名的注册人、注册域名的个体或者企业名称、域名的注册商、注册的域名商家、域名的注册时间、注册域名的日期、域名到期时间、域名到期的日期、域名的DNS服务器、用来解析域名的DNS服务器、域名的状态、域名解析的状态信息、域名的联系邮箱、域名联系人的邮箱、联系电话域名联系人的电话。其中对域名的注册时间、注册域名的日期、域名到期时间和域名到期的日期识别出年、月、日、时、分、秒六个特征,其他特征使用其相应的个数进行表示。

[0009] 其中DNS数据包中的域名的解析信息包括域名解析的IP信息、域名解析的别名CNAME信息、域名邮件交换服务器地址MX信息、和域名的默认TTL信息,其中域名的默认TTL信息直接使用数值,其余特征使用其相应的个数进行表示。

[0010] 进一步的,步骤S_3具体按以下步骤执行:

$S_{3.1}$:首先进行数据预处理,把获取到的数据分组进行编码,同一组数据当作一个整体样本进行数据处理;

$S_{3.2}$:再进行特征提取,对一组数据中提取DNS多模态特征,对提取的不同类的特征进行超边构造,分别构建决策树进行分类,把决策树分类的叶子节点的所有样本划分到相同的超边之中;

$S_{3.3}$:对不同模态特征构建的超边赋予不同权值,再从超图的矩阵中提取拉普拉斯矩阵并定义如下卷积操作,如式(1);

$$Y = Dv^{-\frac{1}{2}}HWD_e^1H^TD_v^{-\frac{1}{2}}X\Theta \quad \text{式(1)}$$

[0011] 其中,X表示输入节点的特征矩阵,H为超图矩阵,D_v表示顶点度矩阵,D_e表示边

的度矩阵，Θ 表示学习参数；

[0012] $S_{3.4}$：将时序行为特征与超图神经网络进行融合，设计间隔t，每隔时间t使用捕获到的数据节点构建一个超图；设置时间窗口T，T中包含多个t；判断在时间窗口T中的前后超图中是否存在相同节点，存在则对这些节点建立LSTM模型，建立LSTM模型首先在每个时间窗口的超图进行自身的节点特征学习，再对节点进行LSTM模型的时间特征维度更新。

[0013] 建立LSTM模型时，从预处理数据中采样时间窗口为1分钟，时间总长度为1小时的数据进行训练，每个时间窗口采样5N个数据，构建大小为N的超图，在5N个数据中采用5折交叉验证方法，其中4N数据作为训练集，剩余1N作为验证集；其中，采取N为10000的样本进行模型训练。

[0014] 模型训练完成后在线部署使用，当产生新数据时，进行实时检测恶意域名。

[0015] 与现有技术相比，本发明的有益效果包括：

1、结合了域名的多维特征对恶意行为进行刻画，并使用了高阶关联关系的超图模型进行检测，同时对模型进行有效更新，提高了恶意域名的检测效率和准确率。

[0016] 2、通过对已有域名的相似性归纳分析并结合域名时序访问特征来发现潜在在互联网域名访问流量中的未知恶意域名，达到尽早发现网络入侵攻击的效果。

附图说明

[0017] 图1为本发明的方法流程图；

图2为本发明的超图神经网络训练示意图；

图3为本发明的通过相同节点建立节点建立LSTM模型的示意图。

具体实施方式

[0018] 下面结合附图对本发明作进一步详细描述：

参考图1-3所示，本发明提供的基于时序超图神经网络的未知恶意域名检测方法包括以下步骤，如图1所示：

S_1：对未知域名提取其字符特征、解析特征和注册特征，构建多模态域名的特征数据；网络通信对域名进行解析，解析行为可以反应攻击特征。

[0019] S_2：使用决策树方法利用多模态数据构建不同的超边，再对超边构成的超图进行卷积操作，对同一个域名在一段时间前后存在的多个DNS请求，多个域名一起请求，即同一个域名存在时间前后的相关性，不同域名存在相同时间访问的相关性，域名访问频率的时序特征进行卷积操作；

S_3：对获取到的信息进超图神经网络训练，将多个相似域名连接在一起构建超边，超边结合在一起形成超图，再使用超图神经网络进行卷积操作特征聚合，最后通过神经网络的分类器进行节点分类，并建立LSTM模型；

S_4：通过LSTM模型对域名的动静态解析特征进行学习，并对未知恶意域名的检测；

S_5：进行检测模型更新，在进行检测模型更新时，对模型的超图部分使用训练好的超图部分的权重，对模型的分类器部分进行重新训练，得到新的训练模型并进行在线部署使用。

[0020] 本实施例中,在步骤S_1中,由于攻击者使用程序来生成大量的恶意域名,通过域名字符串可以反应此类特征,特征数据包括域名字符串长度、域名字符串的个数、域名中数字占比和域名中数字个数除以域名字符串长度域名的信息熵;和用域名中每个字符出现的频率计算出的熵值。

[0021] 本实施例中,DNS数据包中的域名的注册信息,攻击者为实现攻击通常会注册一批相似的域名来进行IP解析,其中DNS请求信息包括域名的注册人、注册域名的个体或者企业名称、域名的注册商、注册的域名商家、域名的注册时间、注册域名的日期、域名到期时间、域名到期的日期、域名的DNS服务器、用来解析域名的DNS服务器、域名的状态、域名解析的状态信息、域名的联系邮箱、域名联系人的邮箱、联系电话域名联系人的电话。其中对域名的注册时间、注册域名的日期、域名到期时间和域名到期的日期识别出年、月、日、时、分、秒六个特征,其他特征使用其相应的个数进行表示。

[0022] 其中DNS数据包中的域名的解析信息包括域名解析的IP信息、域名解析的别名CNAME信息、域名邮件交换服务器地址MX信息、和域名的默认TTL信息,其中域名的默认TTL信息直接使用数值,其余特征使用其相应的个数进行表示。

[0023] 进一步的,如图2,步骤S_3具体按以下步骤执行:

$S_{3.1}$:首先进行数据预处理,把获取到的数据分组进行编码,同一组数据当作一个整体样本进行数据处理;

$S_{3.2}$:再进行特征提取,对一组数据中提取DNS多模态特征,对提取的不同类的特征进行超边构造,分别构建决策树进行分类,把决策树分类的叶子节点的所有样本划分到相同的超边之中;

$S_{3.3}$:对不同模态特征构建的超边赋予不同权值,再从超图的矩阵中提取拉普拉斯矩阵并定义如下卷积操作,如式(1);

$$Y = Dv^{-\frac{1}{2}}HWD_e^1H^TD_v^{-\frac{1}{2}}X\Theta \text{ 式(1)}$$

[0024] 其中,X表示输入节点的特征矩阵,H为超图矩阵,D_v表示顶点度矩阵,D_e表示边的度矩阵,Θ表示学习参数;

[0025] $S_{3.4}$:域名在某个时间点出现的时候,聚合同一时间类似域名。之后随着时间推移,把自身学习到的历史特征传递下去。本发明将时序行为特征与超图神经网络进行融合,设计间隔t,每隔时间t使用捕获到的数据节点构建一个超图;设置时间窗口T,T中包含多个t;判断在时间窗口T中的前后超图中是否存在相同节点,存在则对这些节点建立LSTM模型,建立LSTM模型首先在每个时间窗口的超图进行自身的节点特征学习,再对节点进行LSTM模型的时间特征维度更新。

[0026] 如图3,建立LSTM模型时,从预处理数据中采样时间窗口为1分钟,时间总长度为1小时的数据进行训练,每个时间窗口采样5N个数据,构建大小为N的超图,在5N个数据中采用5折交叉验证方法, 其中4N数据作为训练集,剩余1N作为验证集;其中,采取N为10000的样本进行模型训练。

[0027] 模型训练完成后在线部署使用,当产生新数据时,进行实时检测恶意域名。

[0028] 恶意域名在语义特征,注册特征,解析特征等静态解析行为上与良性域名具有区分度,同时恶意域名在访问时间的动态解析特征上与良性域名不同。由于超图以及LSTM时

序模型对域名的动静态解析特征进行学习,所以通过本发明训练的模型具有对未知恶意域名的检测能力。

[0029] 上述技术方案只是本发明的一种实施方式,对于本领域内的技术人员而言,在本发明公开了原理的基础上,很容易做出各种类型的改进或变形,而不仅限于本发明上述具体实施例所描述的技术方案,因此前面描述的只是优选的,而并不具有限制性的意义。

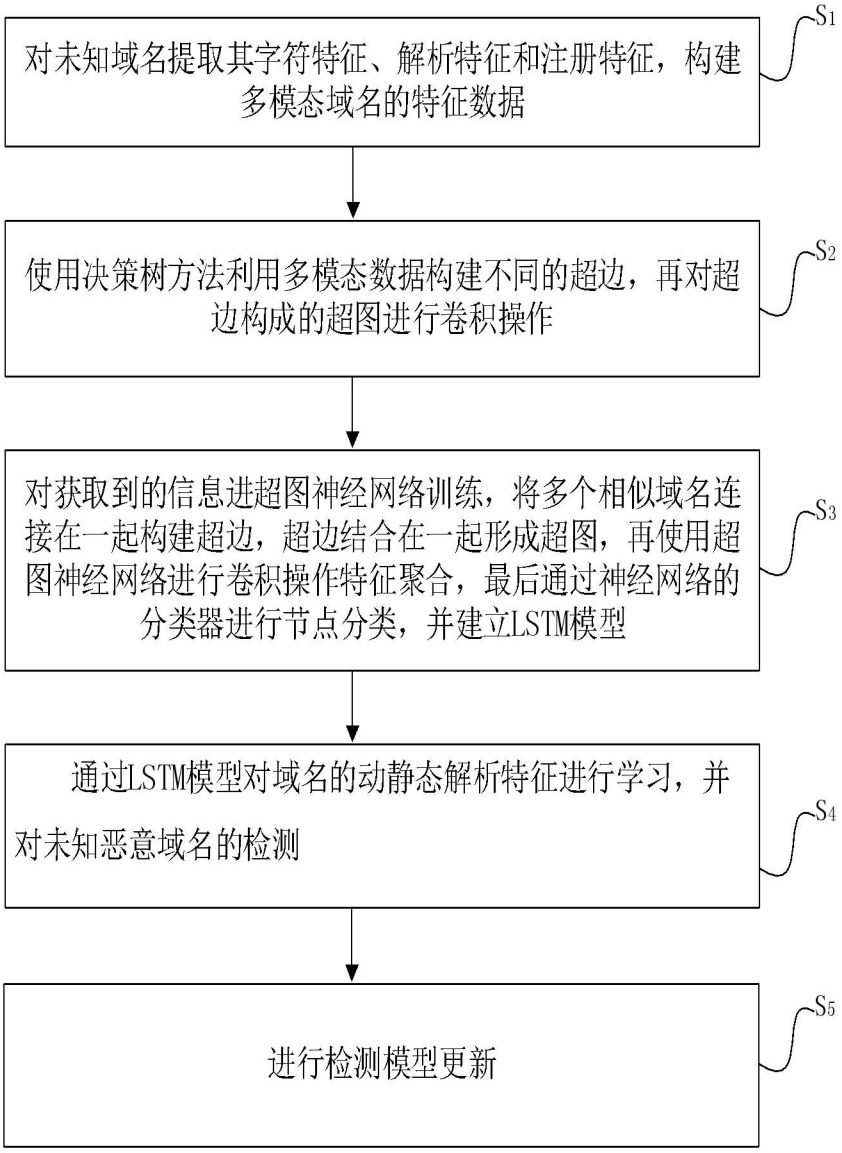

对未知域名提取其字符特征、解析特征和注册特征，构建多模态域名的特征数据
S1
使用决策树方法利用多模态数据构建不同的超边，再对超边构成的超图进行卷积操作
S2
对获取到的信息进超图神经网络训练，将多个相似域名连接在一起构建超边，超边结合在一起形成超图，再使用超图神经网络进行卷积操作特征聚合，最后通过神经网络的分类器进行节点分类，并建立LSTM模型
S3
通过LSTM模型对域名的动静态解析特征进行学习，并对未知恶意域名的检测
S4
进行检测模型更新
S5

图1

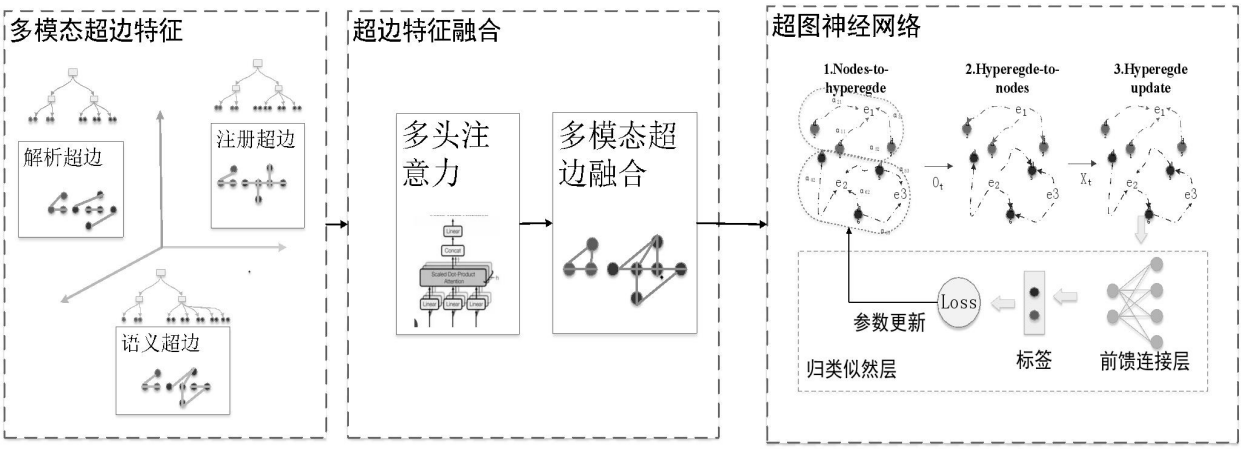

多模态超边特征
解析超边
注册超边
语义超边
超边特征融合
多头注意力
多模态超边融合
超图神经网络
1.Nodes-to-hyperegde
2.Hyperegde-to-nodes
3.Hyperegde update
参数更新
Loss
归类似然层
标签
前馈连接层

图2

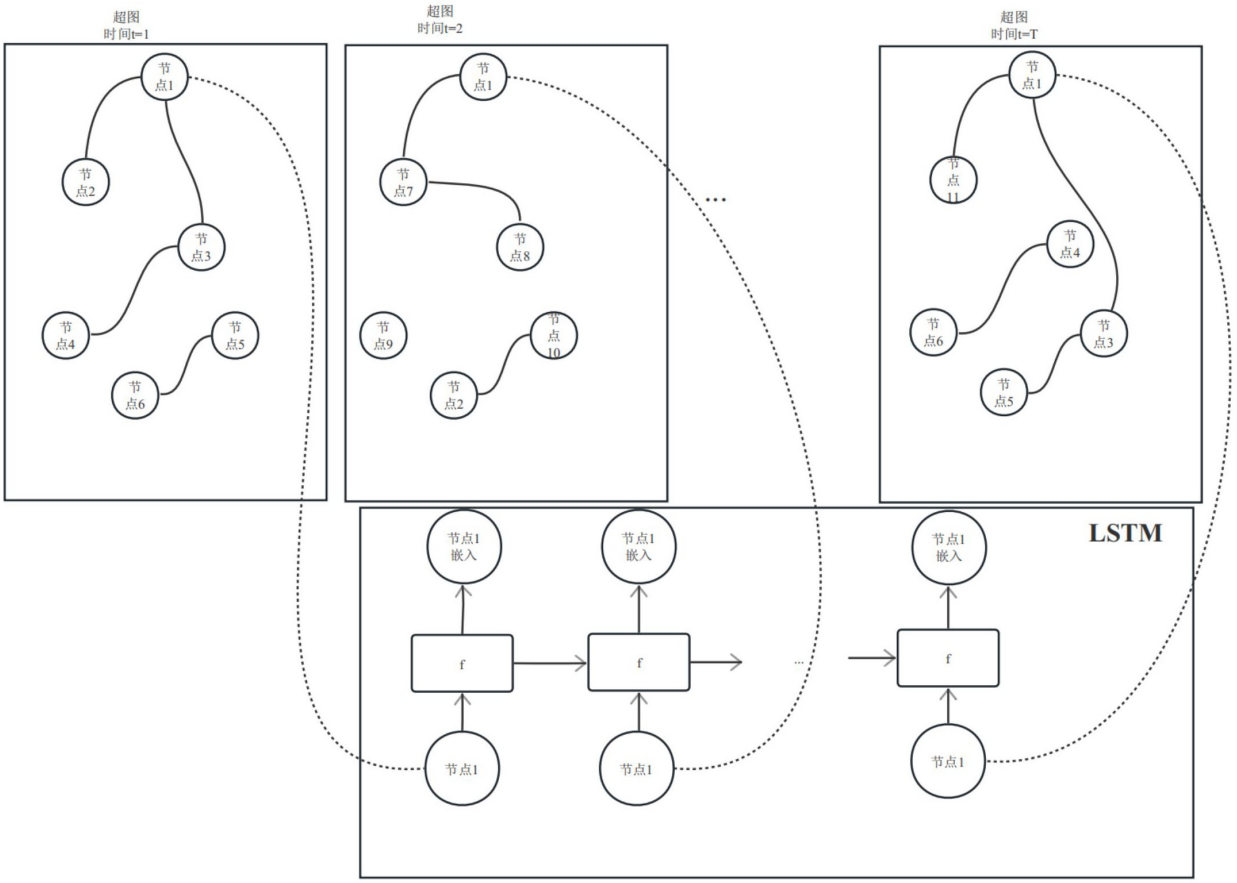

超图
时间t=1
超图
时间t=2
超图
时间t=T
节点1
节点2
节点3
节点4
节点5
节点6
节点7
节点8
节点9
节点10
节点11
LSTM
节点1嵌入
f
节点1

图3