实体消歧的生成事件模式归纳

Kiem Hieu Nguyen[1, 2] 泽维尔·坦尼尔[3, 1] 奥利维尔·雪貂[2] 罗马·贝桑松[2]
(1) 有限责任公司
(2) CEA，清单，Contenus 视觉与工程实验室，F-91191，伊夫特河畔吉夫
(3) 大学巴黎南部
{阮，XTANNIER } Limsi.Fr，{奥利维尔.费雷特，RoMARIC .贝桑松} CEA.FR

摘要

本文提出了事件模式归纳的生成模型。文献中的先前方法仅使用单词来表示实体。但是，除单词以外的元素都包含有用的信息。例如，武装人员比人更具歧视性。我们的模型考虑了此信息，并使用概率主题分布精确地表示了它。我们说明了此类信息在参数估计中起着重要作用。通常，它使主题分布更加连贯和更具区分性。基准数据集上的实验结果凭经验证实了这一增强。

1 介绍

信息提取最初是由 MUC 评估定义的（并且现在仍然定义）（Grishman 和 Sundheim，1996），更具体地说是由模板填充的任务定义的。该任务的目的是将事件角色分配给各个文本提及。模板定义了特定类型的事件（例如地震），与实体持有的语义角色（或时段）相关联（针对地震，其位置，日期，大小及其所造成的损害（Jean-Louis 等，2011））。

*模*式归纳是在没有来自未标记文本的监督的情况下学习这些模板的任务。我们将重点放在事件模式归纳上，并继续为该任务提出的生成模型的趋势。这个想法是基于事件模板中与相同角色对应的实体的相似性，将这些实体分组在一起。例如，在有关恐怖袭击的语料库中，可以将作为杀死，攻击动词的对象的实体组合在一起，并以角色来表征名为 VICTIM。此标识操作的输出是一组群集，这些群集的成员既是单词又是关系，并与它们的概率相关联（请参见后面的图 4 中的示例）。这些集群未标记，但每个集群都代表一个事件槽。

我们在这里的方法是通过消除实体歧义来改善这个最初的想法。某些模棱两可的实体（例如人或士兵）可以匹配两个不同的位置（受害者或作案者）。如果有文章提到恐怖分子已被警察杀害（因此成为杀害对象），则可以将诸如恐怖分子之类的实体与受害者混为一谈。我们的假设是实体的直接上下文有助于消除它们的歧义。例如，人与武装，危险，英雄或无辜相关联的事实可以导致更好的角色归属和定义。然后，我们通过句法关系在模型中介绍实体及其属性之间的关系。

文档级别通常是主题建模的中心概念，但在我们的生成模型中并未使用。这样就形成了一个更简单，更直观的模型，其中从槽生成观察值，该槽由实体，谓词和句法属性上的概率分布定义。该模型为进一步扩展提供了空间，因为可以用相同的方式表示实体上的多个观测值。

模型参数通过吉布斯采样估计。我们通过与系统以前的工作类似的方式，通过系统中的插槽和引用中的插槽之间的自动和经验映射来评估此方法的性能。

本文的其余部分安排如下：第 2 节简要介绍了以前的工作；第 2 节简要介绍了以前的工作。在第 3 节中，我们详细介绍了实体和关系表示形式；我们将在第 4 节中描述生成模型，然后在第 5 节中介绍我们的实验和评估。

计算语言学协会第 53 届年会暨第七届自然语言处理国际联合会议论文集，第 188-197 页，中国北京，2015 年 7 月 26-31 日。Qc 2015 计算语言学协会

2 相关工作

尽管已尽力使模板填充尽可能通用，但在很大程度上仍取决于事件的类型。将通用流程与数量有限的领域特定规则（Freedman 等人，2011）或示例（Grishman and He，2014）混合在一起，是一种减少使系统适应另一个领域所需的工作量的方法。按需信息提取（长谷川等，2004； Sekine，2006）和先发信息提取（Shinyama 和 Sekine，2006）的方法试图通过利用从查询中选择的代表性文档中引入的模板，以另一种方式克服这一困难。

事件模式归纳植根于从知识结构的文本中获取的工作，例如早期文本理解系统（DeJong，1982）使用的记忆组织数据包（Schank，1980），以及最近由 Ferret 和 Grau（1997）使用的知识结构。在信息提取（Collier，1998），自动汇总（Harabagiu，2004）和事件 QuestionAnswering（Filatova 等，2006； Filatova，2008）领域已经进行了将这种过程应用于模式归纳的首次尝试。最近，Hasegawa 等人（2004 年）之后的工作已经发展成为信息抽取的弱监督形式，其目标包括模式归纳。然而，它们在实践中主要应用于二元关系提取（Eichler 等，2008； Rosenfeld 和 Feldman，2007； Min 等，2012）。同时，提出了几种在现有框架中专门执行模式归纳的方法：从句图聚类（Qiu 等人，2008），事件序列比对（Reg-neri 等人，2010）或基于 FrameNet 的基于 LDA 的方法。

类语义框架（Bejan，2008 年）。钱伯斯（Chambers）（2013）和张（Cheung）等人提出了更多事件特定的生成模型。（2013）。

最后，由 Balasubramanian 等人改进了 Chambers and Jurafsky（2008），Chambers and Jurafsky（2009），Chambers and Jurafsky（2011）。（2013）和钱伯斯（2013）特别关注事件角色的归纳和事件链的识别，以便通过利用共指解析或事件的时间顺序从文本构建表示形式。所有这些工作也都与从文本归纳脚本的工作有关，或多或少与事件，例如（Frermann 等，2014），（Pichotta 和 Mooney，2014）或（Modi 和 Titov，2014）。

	属性	头	扳机
#1	[武装：恶魔]	男人	[攻击：杀死：
#2	[警察：nn]	站	[攻击：dobj]
#3	[]	警察人	[kill：
#4	[天真：恶毒，年轻：amod]		dobj][伤

图 1：实体表示为（[attributes]，head，[triggers]）的元组。

我们在本文中介绍的工作与钱伯斯（Chambers，2013）一致，我们将在第 5 节中更详细地介绍该工作，并进行定量和定性比较。

3 实体表示

一个实体表示为一个三元组，它包含：单词 h，属性关系列表 A 和触发器关系列表 T。考虑以下示例：

（1）两名武装人员袭击了警察局并杀死了一名警察。一个无辜的年轻人也受伤了。

如图 1 所示，从上面的文本生成了四个等效于四个分离的三元组的实体。标题词是从名词短语中提取的。触发关系由谓词（攻击，杀死，伤口）和依赖类型（对象，对象）组成。属性关系由参数（武装，警察，年轻人）和依存关系类型（形容词，名词或言语修饰语）组成。在与触发器的关系中，主词是自变量，在与属性的关系中，谓词是谓词。我们使用 Stanford NLP 工具包（Manning 等人，2014）进行解析和共引用解析。

如果单词是名词或专有名词，并且与至少一个触发器相关，则将其提取出来；代词被省略。如果单词是动词或主观名词，并且该单词用作其主语，宾语或介词，则提取该单词的触发条件。我们在 WordNet 中使用类别名词.EVENT 和名词.ACT 作为主语列表。一个主词可以有多个触发器。这些多重关系可以来自单个句子内部的句法协调，例如在说明示例的第一句中就是这种情况。它们也可以代表共同引用

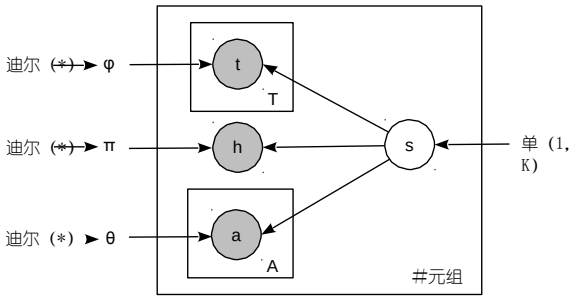


图 2：事件感应的生成模型。

跨句子链接，因为我们使用共指解析将合并提及的触发器合并到文档中的同一实体。共指是事件归纳的有用来源（Chambers 和 Jurafsky，2011； Chambers，2013）。最后，如果属性是形容词，名词或

一个动词，用作形容词，言语或名词
头词的 inal 修饰词。如果有多个修饰词，则仅选择最接近主词的形式。这种“最佳选择”试探法允许省略实体的非歧视属性。

4 生成模型

4.1 型号说明

图 2 显示了我们模型的板符号。对于代表实体 e 的每个三元组，模型首先从统一分布 uni（1，K）中为实体分配时隙 s。然后根据多项式分布 π_s 生成其首字母 h。事件触发关系 Te 的每个 ti 是由多项式分布 φ_s 生成的。属性关系 Ae 的每个 aj 类似地从多项式分布 θ_s 生成。分别从 Dirichlet 先验 dir（α），dir（β）和 dir（γ）生成分布 θ，π 和 φ。

给定一组实体 E，我们的模型（π，φ，θ）定义为

$$P_{\pi,\varphi,\theta}(E) = \prod_{e \in E} P_{\pi,\varphi,\theta}(e) \quad (2)$$

其中每个实体 e 的概率由

$$P_{\pi,\varphi,\theta}(\mathrm{e}) = \mathrm{P}(\mathrm{s}) \times P(h|s) \times \prod_{\text{泰特} \in} P(t|s) \times \prod_{} P(a|s) \quad (3)$$

产生的故事如下：

对于插槽 s←1 至 K
　根据 Dirichlet 先验 dir（α）生成属性分布 θ_s；
　根据 Dirichlet 先验 dir（β）生成水头分布 π_s；
　根据 Dirichlet 先验 dir（γ）生成触发分布 φ_s；
结束
对于实体 e∈E 做
　根据均匀分布生成广告位 *单（1，K）；*
　从多项式分布生成水头 h *；*
　因为我←1 到| Te |做
　　从多项式生成触发器 ti 分布 φ_s；
　结束
　为 j←1 至| Ae |做
　　从多项式生成属性 aj 分布 φ_s；
　结束
结束

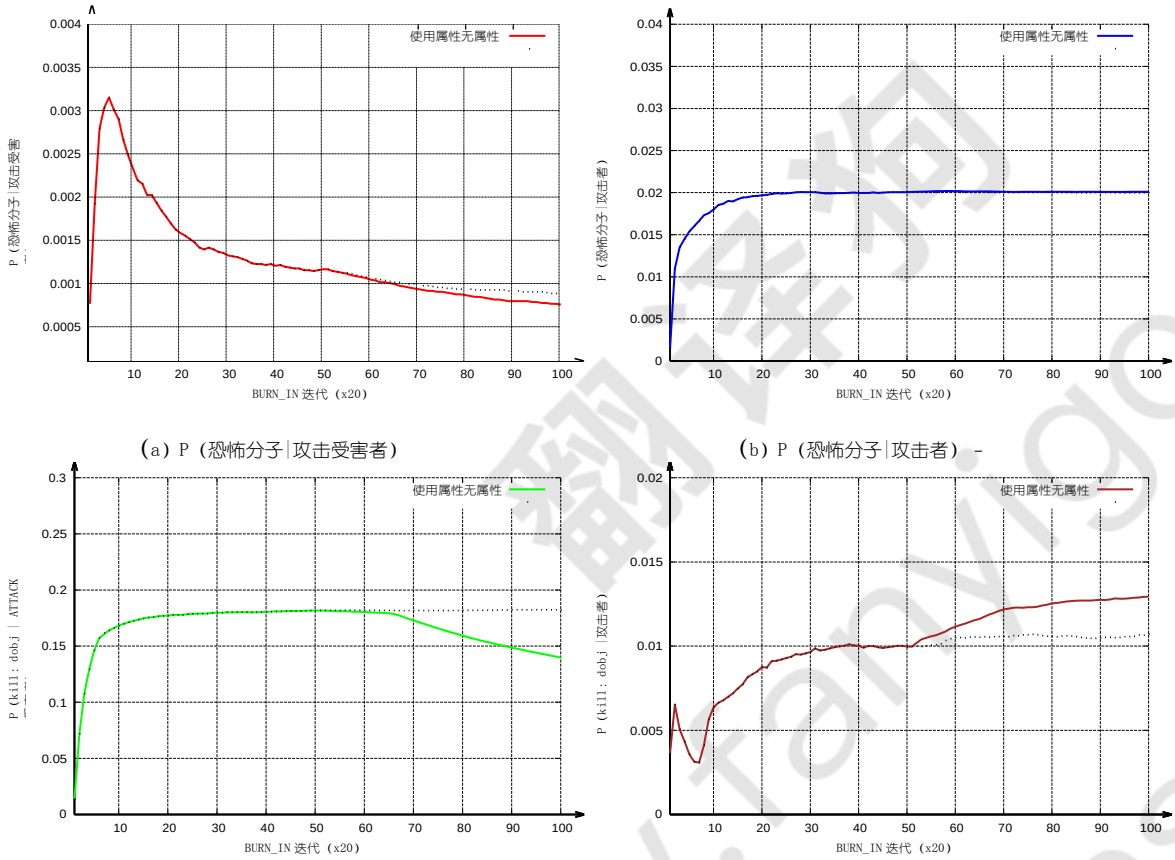
4.2 参数估计

对于参数估计，我们使用 Gibbs 采样方法（Griffiths，2002）。时隙变量 s 通过积分所有其他变量进行采样。

以前的模型（Cheung 等人，2013；Chambers，2013）基于文档级主题建模，该模型源自诸如潜在狄利克雷分配（Blei 等人，2003）之类的模型。相反，我们的模型独立于文档上下文。它的输入是实体三元组的序列。文档边界仅用于过滤的后处理步骤（有关更多详细信息，请参见第 5.3 节）。有一个通用的插槽分布，而不是一个文档的每个插槽分布。此外，通过使用均匀分布作为分类概率的特殊情况，可以忽略时隙先验。基于采样的时隙分配可能取决于初始状态和随机种子。在实施 Gibbs 采样时，我们使用了 10,000 次迭代中的 2,000 次老化。老化的目的是在估计概率分布之前确保参数收敛到稳定状态。而且，在连续样本之间应用 100 的间隔步长，以避免太强的相干性。

特别是为了跟踪概率变化

由于属性关系而产生的联系，我们在第一阶段仅通过头部和触发器关系就进行了特定的老化测试。当时的稳定状态是
用作第二次老化的初始化



(a) P（恐怖分子|攻击受害者）

(b) P（恐怖分子|攻击者）

(c) *P（杀死：dobj | ATTACK 受害者）*

(d) *P（杀戮：DOBJ 攻击 PERP）*

图 3：在采样中使用属性时的概率收敛。属性的使用始于第 50 点（即预烧阶段的 50%）。虚线表示没有属性的收敛；实线显示了属性的收敛。

一起使用了哪些属性，头部和触发器。这个特定的实验设置使我们了解了属性如何修改分布。我们观察到非歧义词或关系（即爆炸，谋杀：nsubj）仅稍作修改，而歧义词（例如人，士兵）或触发器（例如 kill：dobj 或 attack：nsubj）的概率会平稳收敛到另一个稳定状态，在语义上更加连贯。例如，模型有趣地意识到，即使恐怖分子被杀（例如被警察杀死），他实际上并不是袭击的真正受害者。图 3 显示给定 ATTACK 受害者和 ATTACK 犯罪者的恐怖分子和 kill：dobj 的概率收敛。

5 评价

为了与相关工作进行比较，我们对消息理解会议（MUC-4）语料库（Sundheim，1991）进行了评估，使用了传统的精度，召回率和 F 评分模板提取指标。

接下来，我们首先介绍 MUC-4 语料库（第 5.1.1 节），我们详细介绍学习的时隙和参考时隙之间的映射技术（5.1.2）以及模型的超参数（5.1.3）。接下来，我们提出第一个实验（第 5.2 节），该实验显示使用属性关系如何改善总体结果。第二个实验（第 5.3 节）研究了文档分类的影响。然后，我们从定量和定性的角度将我们的结果与以前的方法进行比较，尤其是与钱伯斯（Chambers，2013）进行比较（第 5.4 节）。最后，第 5.5 节致力于错误分析，并特别强调误报的来源。

5.1 实验装置

5.1.1 数据集

MUC-4 语料库包含有关在拉丁美洲发生的恐怖主义事件的 1,700 条新闻报道。语料库分为 1,300 个文档

开发集和四个测试集，每个包含 100 个文档。

我们遵循文献中的规则来保证可比较的结果（Patwardhan 和 Riloff，2007； Chambers 和 Jurafsky，2011）。评估的重点是四种模板类型-ARSON，ATTACK，BOMBING，KIDNAPPING-和四种插槽-犯罪者，乐器，目标和受害者。犯罪者从犯罪者个人和犯罪者组织合并而来。系统答案和参考之间的匹配基于词头匹配。单词定义为短语中最右边的单词，如果该短语包含任何单词，则定义为第一个“ of”中最右边的单词。计算召回率时，将忽略可选的模板和插槽。评估中会忽略模板类型：这意味着答案中的 BOMBING 犯罪者可与 ARSON，ATTACK，BOMBING 或 KIDNAPPING 的犯罪者进行比较
参考资料。

5.1.2 插槽对应

该模型学习 K 个插槽，并将文档中的每个实体分配给其中一个学习的插槽。时隙映射包括将每个参考时隙与等效的学习时隙进行匹配。

请注意，在 K 个获悉的时段中，有些不相关，而另一些（有时质量很高）包含不属于参考的实体（时空信息，主角上下文等）。因此，有意义的是拥有比预期的事件时隙更多的学习时隙。

与文献中的先前工作类似，我们实现了自动的经验驱动插槽映射。每个参考时隙都映射到根据 Fscore 指标在模板提取任务中表现最佳的学习时隙。在这里，必须分别映射两个不同模板（例如 ATTACK 受害者和 KIDNAPPING 受害者）的两个相同的插槽。图 4 显示了两个学习到的插槽中最常见的单词，它们被映射到 BOMBING 仪器和 KIDNAPPING 受害者。然后保留此映射以进行测试。

5.1.3 参数调整

我们首先在开发集上调整了模型的超参数。时隙数设置为 K =35。狄利克雷先验设置为 $\alpha = 0.1$，$\beta = 1$ 和 $\gamma = 0.1$。该模型是从整个数据集中学习的。插槽映射是在 tst1 和 tst2 上完成的。来自 tst3 和 tst4 的输出为 eval-

图 4：模型 HT + A 为学习到的插槽映射到 BOMBING 仪器和 KIDNAPPING 受害者所学习的属性，头部和触发器分布。

使用参考进行计算，并在十次运行中求平均值。

5.2 实验 1：使用实体属性

在该实验中，比较了我们模型的两个版本：HT + A 使用实体头，事件触发关系和实体属性关系。HT 仅使用实体头和事件触发器并省略属性。

我们研究了属性关系带来的收益，着重研究了在获得或缺少共指信息时它们之间的关系。模型输入的变体分别命名为 single，multi 和 coref。每个实体的单个输入只有一个事件触发器。像武装人员这样的文字袭击了警察局并杀害了一名警察，导致实体男子的身高提高了三倍：（armed：amod，man，Attack：nsubj）和（armed：amod，man，kill：nsubj）。在多输入中，一个实体可以具有多个事件触发器，导致上面的文本进入三元组（armed：amod，man，[attack：nsubj，kill：nsubj]）。coref 输入比 multi 输入丰富，因为除了来自同一句子的触发器外，链接到相同 corefered 实体的触发器也合并在一起。例如，如果上述示例中的人在 3 小时后与他一起被捕，则合并的三人组变为（武装：amod，人，[攻击：nsubj，kill：nsubj，逮捕：dobj]）。这些模型和数据组合的板符号在图 5 中给出。

表 1 显示了使用带有和不带有共引用的属性时的一致改进。完整模型在具有 coref-f 的输入上获得 40.62 F 分数的最佳性能。

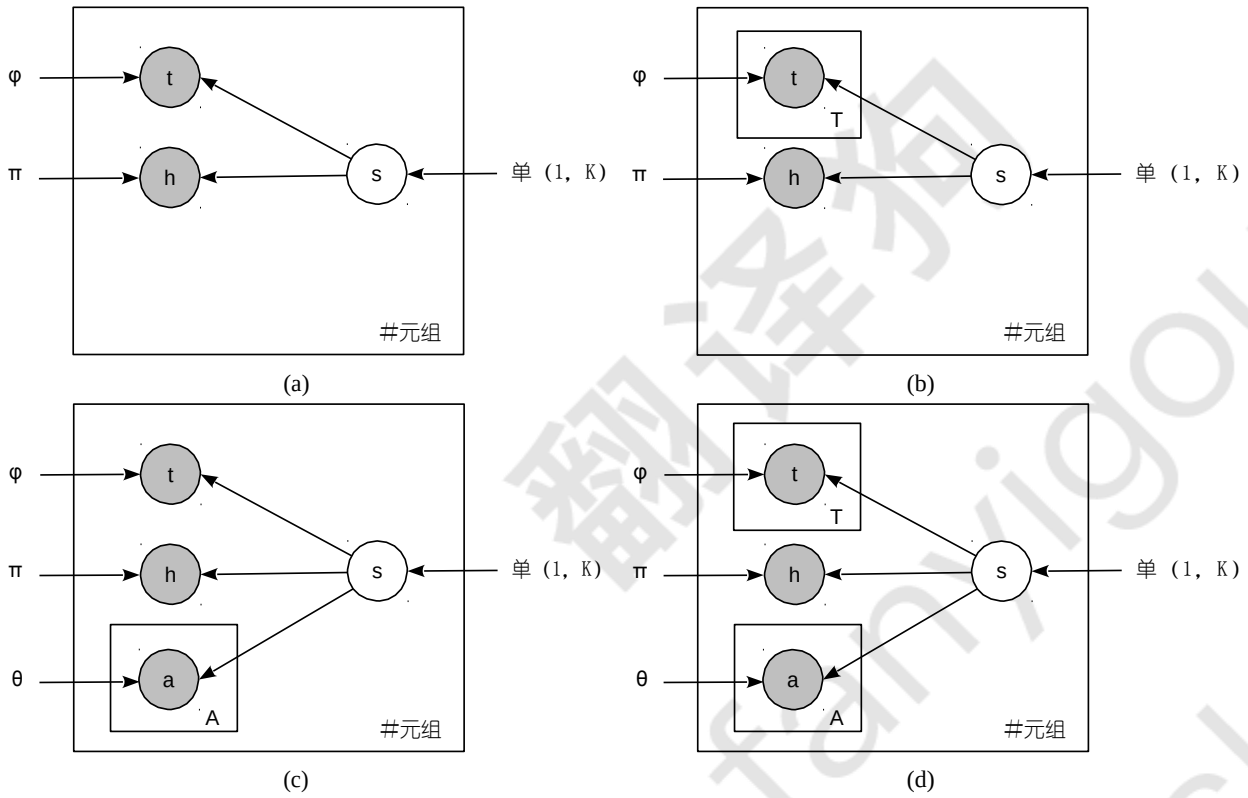


图 5：模型变体（为简单起见，省略了 Dirichlet 先验）：5a）HT 模型基于单个数据运行。该模型等效于 T = 1 的 5b）。5b）HT 模型在多数据上运行；5c）HT + A 模型基于单个数据运行；5d）HT + A 模型在多数据上运行。

数据	HT			ht_a		
	P	R	F	P	R	F
单身	29.59	51.17	37.48	30.22	52.41	38.33
多点	29.32	52.21	37.52	30.82	51.68	38.55
酷睿	39.99	53.53	40.01	32.42	54.59	40.62

表 1：使用属性的改进。

特质。使用模型中的属性和共参考来生成输入数据，将获得 3 个 F 分数点的增益。

5.3 实验 2：文件分类

在第二个实验中，我们评估了模型
以及文档分类的后处理步骤。

MUC-4 语料库包含许多“不相关”的文档。如果文档不包含模板，则不相关。在开发集中的 1,300 个文档中，有 567 个无关紧要。最具挑战性的部分是，有许多恐怖分子
*例*如无关的文件中出现的炸弹，武力，游击队。这使得筛选出这些文档很重要，但很困难。作为文件分类-

我们的模型未明确执行分类，因此需要后处理步骤。预计文档分类会减少无关文档中的误报，而不会显着减少召回率。

给定一个具有插槽分配实体的文档 d 和由插槽映射产生的一组映射的插槽 Sm，我们必须确定该文档是否相关。我们将文档的相关性得分定义为：

$$\text{相关性}(d) = \frac{\sum_{e \in d} \sum_{\text{泰特}} P(T \mid SE)}{\sum_{e \in d} \sum_{\text{泰特}} P(T) \, SE} \quad (4)$$

其中 e 是文档 d 中的实体；se 是分配给 e 的时隙值；t 是触发器 Te 列表中的事件触发器。

等式（4）将实体的分数定义为给定时隙的触发器的条件概率之和。文档的相关性分数与分配给映射的版位的实体的分数成正比。如果此相关性得分高于阈值 λ，则该文档被视为相关。调整 λ = 0.02 的值

系统	P	R	F
ht_a	32.42	54.59	40.62
HT + A +文档分类	35.57	53.89	42.79
HT + A + oracle 分类	44.58	54.59	49.08

表 2：将文档分类作为后处理进行的改进。

通过最大程度地提高文档分类的 F 得分来确定发展。

表 2 显示了应用文档分类时的改进。随着不相关文档的误报被过滤掉，精度会提高。召回损失的原因是相关文档被错误地过滤掉。但是，这种损失并不明显，总体 Fscore 最终增加了 5%。我们还将结果与“ oracle”分类器进行比较，该分类器将删除所有不相关的文档，同时保留所有相关的文档。此 oracle 分类的性能表明，文档分类还有一些进一步改进的空间。

不相关文档过滤是大多数有监督和无监督方法所应用的技术。有监督的方法更喜欢在句子或短语级别进行相关性检测（ Patwardhan 和 Riloff ， 2009 ； Patwardhan 和 Riloff，2007）。至于几种无监督的方法，Chambers（2013）在他的主题模型中包括了文档分类。Chambers and Jurafsky（2011）和 Cheung 等。（2013 年）使用学习的聚类通过从有关事件触发器的事后统计中估计文档相对于模板的相关性，对文档进行分类。

5.4 与最新技术的比较

为了更深入地比较我们的结果与文献中的最新技术。我们重新实现了 Chambers（2013）提出的方法，并将属性分布集成到他的模型中（如图 6 所示）。

此模型与我们的模型之间的主要区别如下：

1. Chambers（2013）的完整模板模型添加了将事件链接到文档的分布 ψ。由于没有理由连接文档和插槽（文档可能包含对多个模板的引用，并且插槽映射不依赖于文档级别），因此这会使模型更加复杂，并且可能不太直观。本文档的好处分发是因为它可以对无关文档进行免费分类，从而避免了分类之前或之后的处理。但是，文档相关性的问题非常针对于 MUC 语料库和评估方法。在更一般的用例中，将没有“不相关”的文档，只有涉及各种主题的文档。

2. 每个实体都链接到事件变量 e。此事件为每个提及的实体生成一个谓词（请注意，提及某个实体就是该实体在文档中（例如在共指链中）所有出现的实体）。相反，我们的工作集中在一个概率模型可以在同一位置具有多个观测值的事实。平等对待多个触发器和多个属性。多个属性和多个触发器的来源不仅来自文档级别的共引用，而且还来自依赖关系（或者甚至来自域级的实体共引用，如果有的话）。因此，可以说我们的模型在建模和输入数据方面可以更好地概括。

3. Chambers（2013）在采样过程中应用了启发式约束，强加相同谓词的主语和宾语（例如 kill: nsubj 和 kill: dobj）没有分布在同一时隙中。我们的模型不需要这种启发式。

系统	P	R	F
Cheung 等。（2013 年）	32	37	34
钱伯斯和尤拉夫斯基（2011）	**48**	25	33
《钱伯斯》（2013 年）（纸	41	41	41
HT + A +文档分类	36	**54**	**43**

表 3：与最新的无监督系统的比较。

钱伯斯（Chambers，2013 年）没有充分说明有关数据预处理和模型参数的一些细节；因此，我们对模型的实施（应用于相同的数据）导致的结果与发布的结果略有不同。这就是为什么我们在这里给出两个结果的原因（表 3 中的纸张值，表 4 中的重新实现值）。

表 3 显示，我们的模型在召回时大大优于其他模型。它实现了

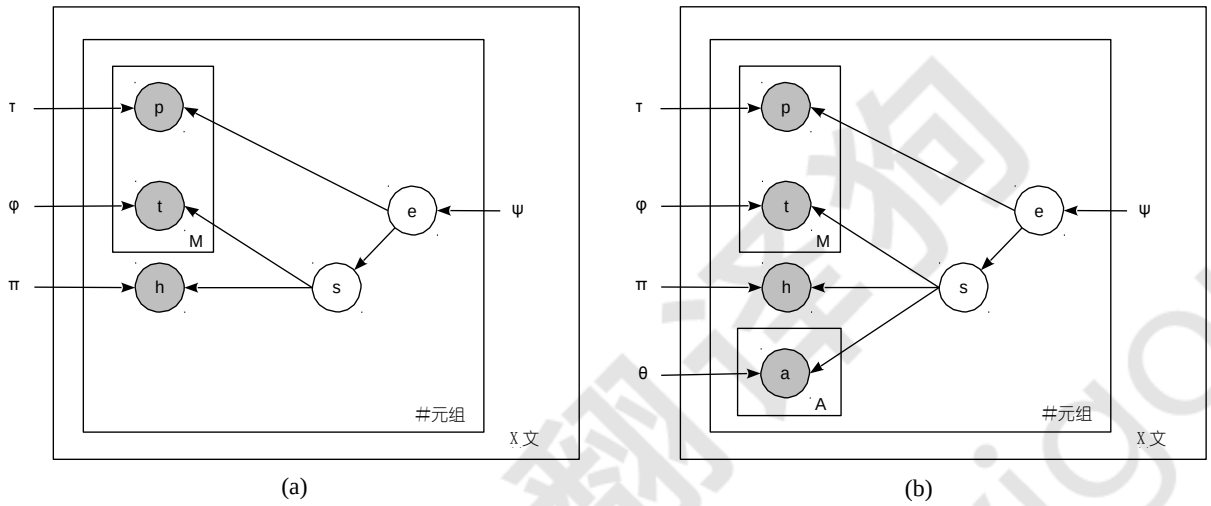


图 6：《钱伯斯的变化》（2013 年）模型：6a）原始模型；6b）原始模型+属性分布。

钱伯斯（2013）	P	R	F
原始 reimpl。	38.65	42.68	40.56
原始 reimpl。+属性	39.25	43.68	41.31

表 4：重新启用分庭的业绩（2013 年）。

最佳整体 F 分数。此外，如我们的实验所述，可以通过更复杂的文档分类进一步提高精度。有趣的是，在 Chambers（2013）提出的模型中，使用属性也被证明是有用的（如表 4 所示）。

5.5 误差分析

我们对 HT + A + doc 的输出执行了错误分析。分类以检测假阳性（FP）的来源。38%的 FP 是参考文献中从未提及的内容。在这 38%的错误中，攻击者和杀手是最常见的错误。这些词可能指的是攻击者。但是，这些引用在参考文献中并未出现，可能是因为人类注释者认为它们过于笼统。除了此类通用术语外，其他分配也是系统的明显错误，例如，窗户，门或墙壁作为物理目标；作为犯罪者的行动或屠杀；爆炸或射击为乐器。这些类型的错误是由于这样一个事实造成的，即在我们的模型中（如钱伯斯（Chambers）（2013）一样），插槽数量是固定的，并不等于参考插槽的实际数量。

另一方面，有 62%的 FP 被提及在参考中至少出现一次的实体。排在首位的是诸如游击队，团体和反叛者的肇事者。如果模型伴随有 announce：nsubj 之类的触发器，则能够将游击队分配给归因槽。但是，描述准恐怖主义事件（例如威胁，威胁，军事冲突）的触发因素也被归类为犯罪者。同样，经常提及的单词，例如炸弹（仪器），建筑物，房屋，办公室（目标），无论它们之间的关系如何，都倾向于系统地归类到这些位置。增加插槽数量（以提高其内容的清晰度）总体上无济于事。这是因为 MUC 语料库非常小，并且倾向于恐怖主义事件。如 Chambers（2013）所述，添加更高级别的模板类型可以部分解决此问题，但可以减少召回率（如表 3 所示）。

6 结论与观点

我们提出了一个生成模型，用于表示事件模板中实体所扮演的角色。我们专注于使用实体的即时上下文，并提出了比先前工作中提出的模型更简单，更有效的模型。我们在 MUC-4 语料库上评估了该模型。

即使我们的结果优于其他非监督方法，我们仍然与监督系统获得的结果相去甚远。可以通过几种方式获得改进。首先，MUC-4 语料库的特征是一个限制因素。语料库很小，并且各个模板的角色相似，无法反映现实。

更大的语料库，甚至被部分注释，但呈现出更好的模板种类，可能导致截然不同的方法。

正如我们所展示的，我们的模型带有所有类型关系的统一表示。这为使用多种类型的关系（句法，语义，主题等）开辟了道路，以完善集群。

最后但并非最不重要的一点是，评估协议已成为一种事实上的标准，它是非常不完善的。最值得注意的是，最终通过参考时隙进行映射的方式可能会对结果产生很大影响。

致谢

这项工作部分由科学合作基金会“ Campus ParisSaclay”（FSC）在 Digiteo ASTRE No. 2013-0774D 项目下资助。

参考文献

Niranjan Balasubramanian, Stephen Soderland, Mausam 和 Oren Etzioni。2013。大规模生成一致性事件模式。在 2013 年自然语言处理的经验方法会议（EMNLP 2013），第 1721–1731 页，美国华盛顿州西雅图，10 月。

考斯敏·阿德里安·贝扬（Cosmin Adrian Bejan）。2008 年。从文本中无监督地发现事件场景。在第二十一届国际佛罗里达人工智能研究协会会议（FLAIRS 2008），第 124-129 页，佛罗里达椰子树。

David M. Blei，Andrew Y. Ng 和 Michael I.Jordan。2003。潜在的 Dirichlet 分配。机器学习研究杂志，3 月 3：993–1022。

纳塔奈尔·钱伯斯（Nathanael Chambers）和丹●尤拉夫斯基（Dan Jurafsky）。2008 年。叙事事件链的无监督学习。在 ACL-08：HLT 中，第 789-797 页，俄亥俄州哥伦布，6 月。

纳塔奈尔·钱伯斯（Nathanael Chambers）和丹●尤拉夫斯基（Dan Jurafsky）。2009 年。叙事图式及其参与者的无监督学习。在 ACL 第 47 届年会和 AFNLP 第四届国际自然语言处理国际联合会议（ACL-IJCNLP'09）的联席会议上，Suntec，第 602-610 页，
8 月，新加坡。

纳塔奈尔·钱伯斯（Nathanael Chambers）和丹●尤拉夫斯基（Dan Jurafsky）。2011。没有模板的基于模板的信息提取。在计算语言学协会：人类语言技术协会（ACL 2011）第 49 届年会上，第 976-986 页，美国俄勒冈州波特兰，6 月。

纳撒尼尔·钱伯斯（Nathanael Chambers）。2013。事件模式归纳与概率实体驱动模型。在 2013 年自然语言处理经验方法会议论文集，第 1797-1807 页，美国华盛顿州西雅图，10 月。

Kit Jackie Chi Cheung，Hoifung Poon 和 Lucy Vanderwende。2013。概率框架归纳。在《计算语言学协会：人类语言技术》 2013 年北美分会会议记录中，第 837-846 页。

R. 科利尔。1998。用于信息提取的自动模板创建。博士论文，谢菲尔德大学。

杰拉尔德·德容。1982 年。FRUMP 系统概述。在 W. Lehnert 和 M. Ringle 中，自然语言处理策略的编辑，第 149-176 页。劳伦斯●埃尔鲍姆协会。

凯瑟琳·艾希勒（Kathrin Eichler），霍尔默●汉森（Holmer Hemsen）和冈特●诺伊曼（GuânterNeumann）。2008。从 Web 文档中无监督的关系提取。在摩洛哥马拉喀什举行的第六届语言资源与评估会议（LREC'08）中。

奥利维尔·弗雷特（Olivier Ferret）和布里吉特●格劳（Brigitte Grau）。1997。建立情节记忆的聚合程序。在第 15 届国际人工智能联合会议（IJCAI-97），第 280-285 页，日本名古屋。

Elena Filatova，Vasileios Hatzivassiloglou 和 Kathleen McKeown。2006。域模板的自动创建。在 21 *st*国际计算语言学会议和第 44 届计算语言学协会年度会议（COLING-ACL 2006），第 207-214 页，澳大利亚悉尼。

埃琳娜●菲拉托娃（Elena Filatova）。2008。针对事件的问题解答和领域建模的无监督关系学习。博士论文，哥伦比亚大学。

Marjorie Freedman，Lance Ramshaw，Elizabeth Boschee，Ryan Gabbard，Gary Kratkiewicz，Nicolas Ward 和 Ralph Weischedel。2011。极端提取 -一周内的机器阅读。在 2011 年自然语言处理的经验方法会议（EMNLP 2011）上，第 1437–1446 页，英国苏格兰爱丁堡，7 月。

Lea Frermann，Ivan Titov 和 Manfred Pinkal。2014 年。无监督脚本知识归纳的多层贝叶斯模型。在计算语言学协会欧洲分会第 14 届会议（EACL 2014），第 49-57 页，瑞典哥德堡，4 月。

汤姆●格里菲斯（Tom Griffiths）。2002。Gibbs 抽样在潜在 Dirichlet 分配的生成模型中进行。斯坦福大学技术报告。

拉尔夫格里什曼和何一凡。2014。信息提取定制程序。在 Petr Sojka，Ale Hork，Ivan Kopeek 和 Karel Pala 的编辑中，第 17 届国际文本，语音和对话会议（TSD 2014），计算机科学讲义第 8655 卷，第 3-10 页。施普林格国际出版社。

拉尔夫·格里什曼（Ralph Grishman）和贝丝●桑德海姆（Beth Sundheim）。1996 年。消息理解会议 6：简要历史。在第 16 届国际计算语言学会议（COLING'96），第 466-471 页，丹麦哥本哈根。

散*打（Sanda Harabagiu）。2004。增量主题表示法。8 月在瑞士日内瓦举行的第 20 届国际计算语言学会议论文集（COL-04）中。*

长谷川孝明（Takaaki Hasegawa），中本聪（Satoshi Sekine）和拉尔夫●格里什曼（Ralph Grishman）。2004。发现大型语料库中命名实体之间的关系。在计算语言学协会会议（ACL'04）的 42 *较*会议上，第 415-422 页，西班牙巴塞罗那。

Ludovic Jean-Louis，Romaric Besanon 和 Olivier Ferret。2011。信息提取中基于文本分割和基于图的模板填充方法。在第五届国际自然语言处理联合会议（IJCNLP 2011），第 723-731 页，泰国清迈。

Christopher D.Manning，Mihai Surdeanu，John Bauer，Jenny Finkel，Steven J.Bethard 和 David McClosky。2014 年。斯坦福大学 CoreNLP 自然语言处理工具包。在《计算语言学协会第 52 届年会论文集：系统论证》，第 55-60 页，美国巴尔的摩，6 月。

博南敏，史淑明，拉尔夫·格里什曼和林紫玉。2012。用于大规模无监督关系提取的集成语义。在 2012 年自然语言处理和计算自然语言学习经验方法联合会议（EMNLP-CoNLL 2012，第 1027-1037 页），韩国济州岛。

Ashutosh Modi 和 Ivan Titov。2014。诱导脚本知识的神经模型。在第 18 届计算自然语言学习会议（CoNLL 2014），第 49-57 页，密歇根州安阿伯。

Siddharth Patwardhan 和 Ellen Riloff。2007。有效的信息提取与语义亲和力模式和相关区域。在 2007 年自然语言处理和计算自然语言学习经验方法联合会议（EMNLP-CoNLL 2007）的会议记录中，第 717-727 页，六月，捷克共和国布拉格。

Siddharth Patwardhan 和 Ellen Riloff。2009。信息抽取的短语和句子证据统一模型。在 2009 年会议记录中

自然语言处理中的经验方法会议（EMNLP 2009），第 151-160 页。

Karl Pichotta 和 Raymond Mooney。2014 年。多参数事件的统计脚本学习。在计算语言学协会欧洲分会第 14 届会议（EACL 2014），第 220-229 页，瑞典哥德堡。

龙秋，Kan 敏欣和蔡达成。2008。场景模板创建中的建模上下文。在第三届国际自然语言处理联合会议（IJCNLP 2008），第 157-164，印度海得拉巴。

Michaela Regneri，Alexander Koller 和 Manfred Pinkal。2010。通过网络实验学习脚本知识。在 7 月，计算语言学协会第 48 届年会（ACL 2010），第 979-988 页，瑞典乌普萨拉。

本杰明·罗森菲尔德（Benjamin Rosenfeld）和罗恩●费尔德曼（Ronen Feldman）。2007。无监督关系识别的聚类。在第 16 届 ACM 信息和知识管理会议（CIKM'07）会议上，第 411-418 页，葡萄牙里斯本。

罗杰·C·尚克 1980.语言与记忆。认知科学，4：243–284。

S*ekoshi Sekine。2006。按需信息提取。在 21 圣国际计算语言学会议和第 44 届计算语言学协会年度会议（COLING-ACL 2006），第 731-738 页，澳大利亚悉尼。*

新山雄介和 Sekoshi Sekine。2006 年。使用无限制关系发现的先发信息提取。在 HLT-NAACL 2006 中，第 304-311 页，美国纽约。

贝丝·桑德海姆。1991 年。第三次信息理解评估和会议（MUC-3）：第一阶段状态报告。在语音和自然语言研讨会的论文集中，HLT '91，第 301-305 页。