(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国 际 局

(43) 国际公布日
2023 年 7 月 6 日 (06.07.2023)

WIPO | PCT

(10) 国际公布号
WO 2023/124647 A1

(51) 国际专利分类号:
G06F 16/33 (2019.01)

(21) 国际申请号: PCT/CN2022/133166

(22) 国际申请日: 2022 年 11 月 21 日 (21.11.2022)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202111662331.8 2021年12月30日 (30.12.2021) CN

(71) 申请人: 安徽听见科技有限公司(ANHUI IFLYREC TECHNOLOGY CO., LTD.) [CN/CN]; 中国安徽省合肥市高新区习友路3333号中国声谷A区2号科研楼16-19层, Anhui 230088 (CN)。

(72) 发明人: 闫莉(YAN, Li); 中国安徽省合肥市高新区习友路3333号中国声谷A区2号科研楼16-19层, Anhui 230088 (CN)。 戚婷(QI, Ting); 中国安徽省合肥市高新区习友路3333号中国声谷A区2号科研楼16-19层, Anhui 230088 (CN)。 高建清(GAO, Jianqing); 中国安徽省合肥市高新区习友路3333号中国声谷A区2号科研楼16-19层, Anhui 230088 (CN)。 孙境廷(SUN, Jingting); 中国安徽省合肥市高新区习友路3333号中国声谷A区2号科研楼16-19层, Anhui 230088 (CN)。

(74) 代理人: 北京集佳知识产权代理有限公司(UNITALEN ATTORNEYS AT LAW CO., LTD.); 中国北京市朝阳区建国门外大街22号赛特广场7层, Beijing 100004 (CN)。

WO 2023/124647 A1

(54) **Title:** SUMMARY DETERMINATION METHOD AND RELATED DEVICE THEREOF

(54) 发明名称: 一种纪要确定方法及其相关设备

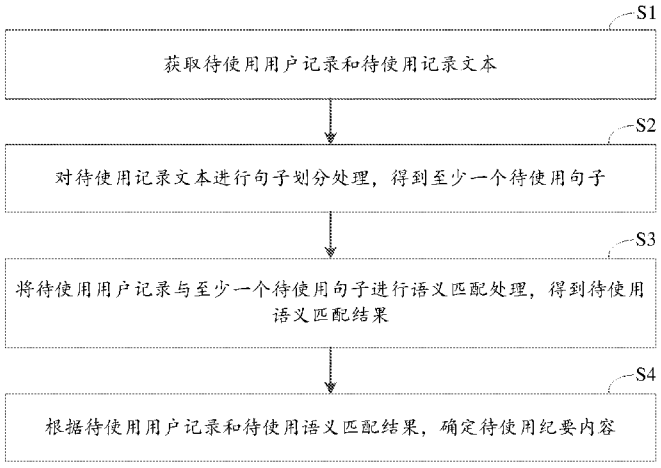


图 1

S1 Perform sentence division processing on record text to be used so as to obtain at least one sentence to be used
S2 Obtain user record to be used and record text to be used
S3 Perform semantic matching processing on user record to be used and sentences to be used so as to obtain semantic matching result to be used
S4 According to user record to be used and semantic matching result to be used, determine summary content

(57) **Abstract:** Provided are a summary determination method and a related device thereof; the method comprises: obtaining a user record to be used and a record text to be used (S1), then performing sentence division processing on the record text to be used so as to obtain at least one sentence to be used (S2); then performing semantic matching processing on the user record to be used and said sentences to be used so as to obtain a semantic matching result to be used (S3); finally, according to the user record to be used and the semantic matching result to be used, determining summary content (S4). According to the semantic matching result to be used, the

[见续页]

(81) 指定国(除另有指明，要求每一种可提供的国家保护)：AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国(除另有指明，要求每一种可提供的地区保护)：ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:
— 包括国际检索报告(条约第21条(3))。

degree of semantic matching between the user record to be used and at least one sentence in the record text to be used can be represented, such that the user can quickly find, on the basis of the "semantic matching result to be used", conference content related to the user record to be used from from the record text to be used, so that the summary determination efficiency can be effectively improved.

(57) 摘要：一种纪要确定方法及其相关设备，该方法包括：在获取到待使用用户记录和待使用记录文本（S1）之后，可以先对该待使用记录文本进行句子划分处理，得到至少一个待使用句子（S2）；再将待使用用户记录与这些待使用句子进行语义匹配处理，得到待使用语义匹配结果（S3）；最后，根据待使用用户记录和待使用语义匹配结果，确定待使用纪要内容（S4）。其中，因上述"待使用语义匹配结果"能够表示出待使用用户记录与待使用记录文本中至少一个句子之间的语义匹配程度，使得用户能够基于该"待使用语义匹配结果"快速地从待使用记录文本中查找到与该待使用用户记录相关的会议内容，如此能够有效地提高纪要确定效率。

一种纪要确定方法及其相关设备

本申请要求于2021年12月30日提交中国专利局、申请号为202111662331.8、申请名称为“一种纪要确定方法及其相关设备”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及自然语言处理技术领域，尤其涉及一种纪要确定方法及其相关设备。

背景技术

在一些应用场景（例如，会议、演讲等）中，有些参与者（例如，记录人）通常会记录一些简洁的信息点来概括发言人的发言内容（例如，内容提要、会议中产生的结论、或者需要跟进的重要事项等），以便这些参与者后续能够基于这些用户记录，编写纪要内容（例如，会议纪要、报告总结等）。

另外，为了更好地编写纪要内容，这些参与者通常需要重新听会议录制音频（或者，演讲录制音频），并在这些用户记录的基础上编写纪要内容。

然而，因这些参与者需要消耗大量的时间来重新听会议录制音频，如此导致纪要内容的确定过程需要消耗较长时间，从而导致纪要确定效率比较低。

发明内容

本申请实施例的主要目的在于提供一种纪要确定方法及其相关设备，能够提高纪要确定效率。

本申请实施例提供了一种纪要确定方法，所述方法包括：获取待使用用户记录和待使用记录文本；对所述待使用记录文本进行句子划分处理，得到至少一个待使用句子；将所述待使用用户记录与所述至少一个待使用句子进行语义匹配处理，得到待使用语义匹配结果；根据所述待使用用户记录和所述待使用语义匹配结果，确定待使用纪要内容。

本申请实施例还提供了一种纪要确定装置，包括：信息获取单元，用于获取待使用用户记录和待使用记录文本；文本划分单元，用于对所述待使用记录文本进行句子划分处理，得到至少一个待使用句子；语义匹配单元，用于将所述待使用用户记录与所述至少一个待使用句子进行语义匹配处理，得到待使用语义匹配结果；纪要确定单元，用于根据所述待使用用户记录和所述待使用语义匹配结果，确定待使用纪要内容。

本申请实施例提供了一种设备，所述设备包括：处理器、存储器、系统总线；所述处理器以及所述存储器通过所述系统总线相连；所述存储器用于存储一个或多个程序，所述一个或多个程序包括指令，所述指令当被所述处理器执行时使所述处理器执行本申请实施例提供的纪要确定方法的任一实施方式。

本申请实施例提供了一种计算机可读存储介质，所述计算机可读存储介质中存储有指令，当所述指令在终端设备上运行时，使得所述终端设备执行本申请实施例提供的纪要确定方法的任一实施方式。

本申请实施例提供了一种计算机程序产品，所述计算机程序产品在终端设备上运行时，使得所述终端设备执行本申请实施例提供的纪要确定方法的任一实施方式。

基于上述技术方案，本申请具有以下有益效果：

本申请提供的技术方案中，在获取到待使用用户记录和待使用记录文本之后，可以先对该待使用记录文本进行句子划分处理，得到至少一个待使用句子；再将待使用用户记录与这些待使用句子进行语义匹配处理，得到待使用语义匹配结果；最后，根据待使用用户记录和待使用语义匹配结果，确定待使用纪要内容。

其中，因上述“待使用语义匹配结果”能够表示出待使用用户记录与待使用记录文本中至少一个句子之间的语义匹配程度，使得用户能够基于该“待使用语义匹配结果”快速地获知该待使用用户记录在待使用记录文本中所处位置，从而使得该用户能够基于该“待使用语义匹配结果”快速地从待使用记录文本中查找到与该待使用用户记录相关的会议内容，如此能够有效地避免因用户通过重听会议录制音频确定与该待使用用户记录相关的会议内容而导致的时长消耗，从而能够有效地提高纪要确定效率。

另外，本申请实施例还可以基于上述“待使用语义匹配结果”确定出至少一个关键补充语句，以使这些关键补充语句能够表示出针对待使用用户记录的语义补充信息，以便在参考这些关键补充语句、这些目标句子、以及待使用用户记录进行待使用纪要内容确定处理时，用户能够从这些信息中获取到更全面的会议重要语义，如此能够实现针对待使用用户记录进行查缺补漏的目的，从而有利于提高纪要确定效果。

此外，本申请实施例还可以从各个关键补充语句中抽取出要点内容，以使这些要点内容能够更简洁地表示出该待使用用户记录遗漏的重要信息，以便用户能够从这些要点内容中更快速地获知该待使用用户记录的要点补充信息，如此有利于提高纪要确定效率。

附图说明

为了更清楚地说明本申请实施例或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图是本申请的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的附图。

图 1 为本申请实施例提供的一种纪要确定方法的流程图；

图 2 为本申请实施例提供的一种提示信息的展示界面的示意图；

图 3 为本申请实施例提供的一种语义匹配模型的工作原理示意图；

图 4 为本申请实施例提供的另一种语义匹配模型的工作原理示意图；

图 5 为本申请实施例提供的一种待使用图的示意图；

图 6 为本申请实施例提供的一种要点抽取模型的工作原理示意图；

图 7 为本申请实施例提供的一种提示信息更新过程的示意图；

图 8 为本申请实施例提供的一种纪要确定装置的结构示意图。

具体实施方式

发明人在针对纪要确定过程的研究中发现，对于纪要编写人员来说，在听取会议的过程中，该纪要编写人员通常会记录一些简洁的信息点来概括发言人的发言内容（例如，内容提要、会议中产生的结论、或者需要跟进的重要事项等）；而且，在会议结束之后，该纪要编写人员可以先参考这些用户记录重新听会议录制音频，确定出一些与这些用户记录相关的会议内容（例如，该用户记录对应的会议发言语句和/或重点补充语句）；再参考这些用户记录及其相关会议内容编写会议纪要，以使该会议纪要能够更准确。然而，因用于记

录长时会议的会议录制音频的音频时长比较长，使得纪要编写人员需要消耗大量时间来获取与这些用户记录相关的会议内容，如此导致会议纪要编写耗时比较长，从而导致纪要确定效率比较低。

基于上述发现，为了解决背景技术部分的技术问题，本申请实施例提供了一种纪要确定方法，该方法包括：在获取到待使用用户记录和待使用记录文本之后，可以先对该待使用记录文本进行句子划分处理，得到至少一个待使用句子；再将待使用用户记录与这些待使用句子进行语义匹配处理，得到待使用语义匹配结果；最后，根据待使用用户记录和待使用语义匹配结果，确定待使用纪要内容。

其中，因上述“待使用语义匹配结果”能够表示出待使用用户记录与待使用记录文本中至少一个句子之间的语义匹配程度，使得用户能够基于该“待使用语义匹配结果”快速地获知该待使用用户记录在待使用记录文本中所处位置，从而使得该用户能够基于该“待使用语义匹配结果”快速地从待使用记录文本中查找到与该待使用用户记录相关的会议内容，如此能够有效地避免因用户通过重听会议录制音频确定与该待使用用户记录相关的会议内容而导致的时长消耗，从而能够有效地提高纪要确定效率。

另外，本申请实施例不限定纪要确定方法的执行主体，例如，本申请实施例提供的纪要确定方法可以应用于终端设备或服务器等数据处理设备。其中，终端设备可以为智能手机、计算机、个人数字助理（Personal Digital Assitant，PDA）或平板电脑等。服务器可以为独立服务器、集群服务器或云服务器。又如，本申请实施例提供的纪要确定方法可以由终端设备或服务器进行配合执行。

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的附图，对本申请实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例是本申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

方法实施例一

参见图 1，该图为本申请实施例提供的一种纪要确定方法的流程图。

本申请实施例提供的纪要确定方法，包括 S1-S4：

S1：获取待使用用户记录和待使用记录文本。

上述“待使用用户记录”用于表示由相关人员（例如，纪要编写人员、会议速记人员等）在听取会议过程中记录的简洁信息（例如，速记内容、笔记内容、或者会议要点记录等）。例如，上述“待使用用户记录”可以包括：监管漏洞，权威不足，效能低下，分散管理效能漏洞明显等内容。

另外，本申请实施例不限定上述“待使用用户记录”，例如，其可以包括一个词、一个短句、或者一个完整语句。

此外，本申请实施例不限定上述“待使用用户记录”的获取过程，例如，其具体可以包括：在获取到用户上传的会议笔记信息之后，先按照预设切分符，对该会议笔记信息进行切分处理，得到至少一个笔记切分单元；再将各笔记切分单元中无意义符号（如，省略号，破折号，或者位于句首的段落编号等）进行删除处理；最后，将各个笔记切分单元均

确定为待使用用户记录，以使该待使用用户记录尽可能地只包括一个句子。

需要说明的是，上述“预设切分符”可以预先设定（例如，由用户根据个人喜好设定）。例如，其可以为常规句子结束性标点(如，句号、问号、或者叹号等)。

为了便于理解上述“待使用用户记录”的获取过程，下面结合示例进行说明。

作为示例，若上述“会议笔记信息”包括：“监管漏洞，权威不足，效能低下，……分散管理效能漏洞明显。建议……取得一致意见，统一期货和各类衍生品规则”，则可以从该“会议笔记信息”中确定出的两个待使用用户记录具体如下：

->待使用用户记录 1：“监管漏洞，权威不足，效能低下，分散管理效能漏洞明显”。

->待使用用户记录 2：“建议取得一致意见，统一期货和各类衍生品规则”。

上述“待使用记录文本”用于表示会议发言内容；而且本申请实施例不限定该“待使用记录文本”的获取过程，例如，“待使用记录文本”的获取过程具体可以包括：在听取会议过程中，针对实时的会议语音进行语音转写处理，得到待使用记录文本。又如，“待使用记录文本”的获取过程可以包括：在会议结束之后，对会议记录视频（或者会议录音音频）进行语音识别处理，得到该待使用记录文本。又如，“待使用记录文本”的获取过程可以包括：在翻译场景中，可以针对会议记录视频进行语音翻译处理，得到该待使用记录文本。

基于上述 S1 的相关内容可知，当一个用户想要针对某个会议编写纪要内容时，可以从该用户针对该会议上传的会议笔记信息中提取待使用用户记录，以便后续能够从该会议的待使用记录文本中定位出该待使用用户记录的文本位置。

S2：对待使用记录文本进行句子划分处理，得到至少一个待使用句子。

本申请实施例中，在获取到待使用记录文本之后，可以按照预设句子结束标记符（例如，句号、问号、或者叹号等），对该待使用记录文本进行句子划分处理，得到该待使用记录文本的句子集合，以使该句子集合包括该待使用记录文本中各个语句；再将句子集合中各个语句均确定为待使用句子。

S3：将待使用用户记录与至少一个待使用句子进行语义匹配处理，得到待使用语义匹配结果。

上述“待使用语义匹配结果”用于记录待使用用户记录与各个待使用句子之间的语义匹配程度（例如，语义匹配得分等），以使该“待使用语义匹配结果”能够表示出该待使用用户记录携带的语义信息是否与各个待使用句子携带的语义信息相同。

另外，本申请实施例不限定 S3 的实施方式，例如，其可以采用现有的或者未来出现的任意一种语义匹配方法进行实施。又如，为了提高语义匹配效果，可以采用**方法实施例二**所示的 S3 的任一实施方式进行实施。

S4：根据待使用用户记录和待使用语义匹配结果，确定待使用纪要内容。

上述“待使用纪要内容”是指由用户基于待使用用户记录编写的纪要内容。

另外，本申请实施例不限定 S4 的实施方式，例如，当上述“待使用语义匹配结果”包括待使用用户记录与各个待使用句子之间的语义匹配得分时，S4 具体可以为：先根据待使用用户记录与各个待使用句子之间的语义匹配得分，从所有待使用句中筛选出语义匹配得分比较高的一个或者多个待使用句子；再将这些语义匹配得分比较高的待使用句子、以及待使用用户记录均显示给用户（如图 2 所示），以使用户能够基于这些显示内容编写待使

用纪要内容，从而使得该用户能够从这些显示内容中快速地查找到其编写纪要所需内容，如此能够有效地避免因用户通过重听会议录制音频确定与该待使用用户记录相关的会议内容而导致的时长消耗，从而能够有效地提高纪要确定效率。

又如，为了进一步提高纪要确定效果，S4 可以采用下文**方法实施例四**、**方法实施例五**、**方法实施例六**、以及**方法实施例七**所示的 S4 的任一实施方式进行实施。

基于上述 S1 至 S4 的相关内容可知，对于本申请实施例提供的纪要确定方法来说，在获取到待使用用户记录和待使用记录文本之后，可以先对该待使用记录文本进行句子划分处理，得到至少一个待使用句子；再将待使用用户记录与这些待使用句子进行语义匹配处理，得到待使用语义匹配结果；最后，根据待使用用户记录和待使用语义匹配结果，确定待使用纪要内容。

其中，因上述“待使用语义匹配结果”能够表示出待使用用户记录与待使用记录文本中至少一个句子之间的语义匹配程度，使得用户能够基于该“待使用语义匹配结果”快速地获知该待使用用户记录在待使用记录文本中所处位置，从而使得该用户能够基于该“待使用语义匹配结果”快速地从待使用记录文本中查找到与该待使用用户记录相关的会议内容，如此能够有效地避免因用户通过重听会议录制音频确定与该待使用用户记录相关的会议内容而导致的时长消耗，从而能够有效地提高纪要确定效率。

方法实施例二

为了进一步提高纪要确定效果，本申请实施例还提供了上文 S3 的另一种可能的实施方式，其具体可以包括：根据待使用用户记录、至少一个待使用句子以及预先构建的语义匹配模型，确定待使用语义匹配结果。为了便于理解，下面结合示例进行说明。

作为示例，待使用语义匹配结果的确定过程，具体可以包括 S31-S33：

S31：根据至少一个待使用句子以及预设句子标识符，确定待使用匹配文本。

上述“预设句子标识符”用于标识句子的起始位置；而且本申请实施例不限定该“预设句子标识符”，例如，其可以为[CLS]。

上述“待使用匹配文本”用于记录至少一个待使用句子；而且本申请实施例不限定该“待使用匹配文本”的确定过程，例如，其具体可以为：先在各个待使用句子的句首拼接预设句子标识符，得到各个拼接后句子；再将所有拼接后句子进行集合处理，得到待使用匹配文本，以使该待使用匹配文本包括所有拼接后句子。

S32：将待使用用户记录以及待使用匹配文本输入预先构建的语义匹配模型，得到该语义匹配模型输出的匹配预测结果。

其中，“匹配预测结果”用于表示待使用用户记录与待使用匹配文本中至少一个句子之间的语义匹配结果。

上述“语义匹配模型”用于针对该语义匹配模型的输入数据进行语义匹配处理；而且该“语义匹配模型”是一种机器学习模型。

另外，本申请实施例不限定上述“语义匹配模型”的网络结构，例如，其具体可以包括：词编码模块、内容交互模块以及匹配预测模块。

为了便于理解上述“语义匹配模型”的工作原理，下面以上述“匹配预测结果”的确

定过程为例进行说明。

作为示例，利用语义匹配模型确定上述“匹配预测结果”的过程，具体可以包括S321-S323：

S321：根据待使用用户记录、待使用匹配文本以及词编码模块，确定用户记录编码特征和至少一个句子编码特征。

上述“词编码模块”用于针对该词编码模块的输入数据进行编码处理；而且本申请实施例不限定该“词编码模块”，例如，如图3或者图4所示，该“词编码模块”可以采用双向编码（Bidirectional Encoder Representations from Transformers，BERT）预训练模型进行实施。

上述“用户记录编码特征”用于表示待使用用户记录携带的字符信息；而且该“用户记录编码特征”是由词编码模块针对待使用用户记录进行编码处理得到的。

第j个句子编码特征用于表示待使用匹配文本中第j个句子携带的字符信息；而且该“第j个句子编码特征”是由词编码模块针对待使用匹配文本中第j个句子进行编码处理得到的。其中，j为正整数，j≤J，J为正整数，J表示待使用匹配文本中句子个数（也就是，上文“待使用句子”的个数）。

需要说明的是，上述“待使用匹配文本中第j个句子”包括第j个待使用句子；而且本申请实施例不限定该“待使用匹配文本中第j个句子”，例如，其具体可以为：由第j个待使用句子以及预设句子标识符进行拼接得到的第j个拼接后句子。

S322：根据用户记录编码特征、至少一个句子编码特征、以及内容交互模块，得到至少一个内容交互特征。其中，“内容交互模块”用于针对该内容交互模块的输入数据进行语义交互处理。

本申请实施例不限定S322的实施方式，为了便于理解，下面结合三种可能的实施方式进行说明。

在第一种可能的实施方式中，S322其具体可以为：将用户记录编码特征以及至少一个句子编码特征输入内容交互模块，得到该内容交互模块输出的至少一个内容交互特征。需要说明的是，S322的第一种可能的实施方式可以采用下文步骤31-步骤32所示的上述“内容交互模块”的工作原理进行实施。

在第二种可能的实施方式中，S322其具体可以包括步骤11-步骤12：

步骤11：根据用户记录编码特征与各句子编码特征之间的相似度表征数据，从至少一个句子编码特征中筛选出满足第二相似条件的至少一个待处理句子特征。

其中，“用户记录编码特征与第j个句子编码特征之间的相似度表征数据”用于表示该用户记录编码特征与该第j个句子编码特征之间的相似程度，以使该“用户记录编码特征与第j个句子编码特征之间的相似度表征数据”能够表示出待使用用户记录与第j个待使用句子之间的语义相似程度。其中，j为正整数，j≤J，J为正整数。

另外，本申请实施例不限定上述“用户记录编码特征与第j个句子编码特征之间的相似度表征数据”的确定过程，例如，其具体可以为：将用户记录编码特征与第j个句子编码特征之间的余弦距离，确定为该“用户记录编码特征与第j个句子编码特征之间的相似度表征数据”。又如，其具体可以为：将用户记录编码特征与第j个句子编码特征之间的欧

式距离，确定为该“用户记录编码特征与第 j 个句子编码特征之间的相似度表征数据”。

上述“第二相似条件”可以预先设定，例如，假设上述“用户记录编码特征与第 j 个句子编码特征之间的相似度表征数据”越大，则表示待使用用户记录携带的语义信息与第 j 个待使用句子携带的语义信息越相似。基于上述假设可知，当将 J 个句子编码特征按照上述“相似度表征数据”从大到小进行排序时，该“第二相似条件”具体可以为：排列序号低于第一序号阈值（例如，2）。其中，“第一序号阈值”可以预先设定。

需要说明的是，对于上述“第一序号阈值”来说，如果该“第一序号阈值”比较小（如，2），则从上述“至少一个句子编码特征”中筛选出的各个待处理句子特征均与待使用用户记录之间的语义相似性比较高，使得这些待处理句子特征所表征的待使用句子能够更好地针对该待使用用户记录进行会议信息补充；另外，将这些待处理句子特征作为提示信息，可以更好地引导语义匹配模型，确定出待使用记录文本中与该待使用用户记录语义相似的语句，如此能够有效地避免从待使用记录文本中，针对该待使用用户记录确定出语义相关性差别比较大的多个语句，如此有利于提高纪要确定效果。

上述“待处理句子特征”是指满足第二相似条件的句子编码特征。

基于上述步骤 11 的相关内容可知，在获取到用户记录编码特征与至少一个句子编码特征之后，可以从这些句子编码特征中挑选出与该用户记录编码特征最相似的多个句子编码特征（如图 3 所示），分别确定为待处理句子特征，以便后续能够从这些待处理句子特征所表征的待使用句子中确定出与待使用用户记录语义最接近的句子。

需要说明的是，对于图 3 来说，“$r_{j-1}=1$”表示第 j-1 个句子编码特征被挑选出（也就是，第 j-1 个句子编码特征满足第二相似条件）；“$r_j=0$”表示第 j 个句子编码特征未挑选出（也就是，第 j 个句子编码特征不满足第二相似条件）；$s_j=\{w_{j,1},\cdots,w_{j,m}\}$表示第 j 个待使用句子，且$w_{j,1}$表示第 j 个待使用句子中第 1 个语义单元（例如，字和/或词），……（以此类推），$w_{j,m}$表示第 j 个待使用句子中第 m 个语义单元；$q=\{w_1,\cdots,w_n\}$表示待使用用户记录，且w_1表示待使用用户记录中第 1 个语义单元，……（以此类推），w_n表示第 j 个待使用句子中第 n 个语义单元；$\hat{s}_j$表示第 j 个句子编码特征；$\hat{s}_{top1}$表示排列序号为 1 的句子编码特征；$\hat{q}$表示用户记录编码特征；q'表示下文“用户记录表征特征”；s_j''表示第 j 个内容交互特征；y_j表示第 j 个匹配预测得分（也就是，待使用用户记录与第 j 个待使用句子之间的匹配预测得分等）。

步骤 12：将用户记录编码特征以及至少一个待处理句子特征输入内容交互模块，得到该内容交互模块输出的至少一个内容交互特征。

需要说明的是，步骤 12 可以采用下文步骤 31-步骤 32 所示的上述“内容交互模块”的工作原理进行实施。

基于上述 S322 的第二种可能的实施方式的相关内容可知，在获取到用户记录编码特征与至少一个句子编码特征之后，可以先从这些句子编码特征中挑选出与该用户记录编码特征比较相似的句子编码特征；再将这些被挑选出的句子编码特征以及该用户记录编码特征

进行语义交互处理，得到至少一个内容交互特征，以使这些内容交互特征能够表示出这些被挑选出的句子编码特征以及该用户记录编码特征之间的语义交互结果，以便后续能够从这些被挑选出的句子编码特征所表征的待使用句子中，确定出与待使用用户记录语义最接近的句子，如此能够有效地减少语义交互对象个数，从而有利于提高语义交互效率，进而有利于提高纪要确定效率。

实际上，对于上述“至少一个待使用句子”来说，某些待使用句子中可能存在指代词，也可能省略一些前文已有信息，故为了更好地提取出这些待使用句子所携带的语义信息。本申请实施例还提供了上述“语义匹配模型”的另一种可能的实施方式，在该实施方式中，该“语义匹配模型”除了包括上述词编码模块、内容交互模块以及匹配预测模块以外，可以还包括：句编码模块（如图 4 所示）。

需要说明的是，对于图 4 来说，$s_j = \{w_{j,1}, \cdots, w_{j,m}\}$表示第 j 个待使用句子；$q = \{w_1, \cdots, w_n\}$表示待使用用户记录；$\hat{s}_j$表示第 j 个句子编码特征；$\hat{q}$表示用户记录编码特征；$q'$表示下文“用户记录表征特征”；$s_j'$表示第 j 个表征特征；$s_j''$表示第 j 个内容交互特征；$y_j$表示第 j 个匹配预测得分(例如，待使用用户记录与第 j 个待使用句子之间的匹配预测得分等)。

为了便于理解上述“句编码模块”的工作原理，下面结合 S322 的第三种可能的实施方式进行说明。

在第三种可能的实施方式中，当上述“语义匹配模型”还包括句编码模块时，S322 其具体可以包括步骤 21-步骤 22：

步骤 21：将至少一个句子编码特征输入句编码模块，得到该句编码模块输出的至少一个句子表征特征。

上述“句编码模块”用于参考一个句子的上下文信息对该句子进行编码处理；而且本申请实施例不限定该“句编码模块”的实施方式，例如，其可以采用任意一种编码器进行实施。需要说明的是，本申请实施例不限定上述“编码器”，例如，其可以包括两层 Transformer 中 Encoder 层。

第 j 个句子表征特征用于表示第 j 个待使用句子携带的语义信息。其中，j 为正整数，j ≤J，J 为正整数。

步骤 22：根据用户记录编码特征、至少一个句子表征特征、以及内容交互模块，得到至少一个内容交互特征。

本申请实施例不限定步骤 22 的实施方式，例如，在第一种可能的实施方式中，如图 4 所示，步骤 22 具体可以包括：将用户记录编码特征以及至少一个句子表征特征输入内容交互模块，得到该内容交互模块输出的至少一个内容交互特征。需要说明的是，步骤 22 的第一种可能的实施方式可以采用下文步骤 31-步骤 32 所示的上述“内容交互模块”的工作原理进行实施。

另外，为了提高语义交互效率，本申请实施例还提供了步骤 22 的第二种可能的实施方式，其具体可以包括步骤 221-步骤 222：

步骤 221：根据用户记录编码特征与各句子表征特征之间的相似度表征数据，从至少一个句子表征特征中筛选出满足第一相似条件的至少一个待参考句子特征。

其中，“用户记录编码特征与第 j 个句子表征特征之间的相似度表征数据”用于表示该用户记录编码特征与该第 j 个句子表征特征之间的相似程度，以使该“用户记录编码特征与第 j 个句子表征特征之间的相似度表征数据”能够表示出待使用用户记录与第 j 个待使用句子之间的语义相似程度。其中，j 为正整数，j≤J，J 为正整数。

另外，本申请实施例不限定上述“用户记录编码特征与第 j 个句子表征特征之间的相似度表征数据”的确定过程，例如，其具体可以为：将用户记录编码特征与第 j 个句子表征特征之间的余弦距离，确定为该“用户记录编码特征与第 j 个句子表征特征之间的相似度表征数据”。又如，其具体可以为：将用户记录编码特征与第 j 个句子表征特征之间的欧式距离，确定为该“用户记录编码特征与第 j 个句子表征特征之间的相似度表征数据”。

上述“第一相似条件”可以预先设定，例如，假设上述“用户记录编码特征与第 j 个句子表征特征之间的相似度表征数据”越大，则表示待使用用户记录携带的语义信息与第 j 个待使用句子携带的语义信息越相似。基于上述假设可知，当将 J 个句子表征特征按照上述“相似度表征数据”从大到小进行排序时，该“第一相似条件”具体可以为：排列序号低于第二序号阈值。

需要说明的是，对于上述“第二序号阈值”来说，如果该“第二序号阈值”比较小（如，2），则从上述“至少一个句子表征特征”中筛选出的各个待参考句子特征均与待使用用户记录之间的语义相似性比较高，使得这些待参考句子特征所表征的待使用句子能够更好地针对该待使用用户记录进行会议信息补充；另外，将这些待参考句子特征作为提示信息，可以更好地引导语义匹配模型，确定出待使用记录文本中与该待使用用户记录语义相似的语句，如此能够有效地避免从待使用记录文本中，针对该待使用用户记录确定出语义相关性差别比较大的多个语句，如此有利于提高纪要确定效果。

上述“待参考句子特征”是指满足第一相似条件的句子表征特征。

基于上述步骤 221 的相关内容可知，在获取到用户记录编码特征与至少一个句子表征特征之后，可以从这些句子表征特征中挑选出与该用户记录编码特征最相似的多个句子表征特征，分别确定为待处理句子特征，以便后续能够从这些待处理句子特征所表征的待使用句子中确定出与待使用用户记录语义最接近的句子。

步骤 222：将用户记录编码特征以及至少一个待参考句子特征输入内容交互模块，得到该内容交互模块输出的至少一个内容交互特征。

需要说明的是，步骤 222 可以采用下文步骤 31-步骤 32 所示的上述“内容交互模块”的工作原理进行实施。

基于上述步骤 221 至步骤 222 的相关内容可知，在获取到用户记录编码特征与至少一个句子表征特征之后，可以先从这些句子表征特征中挑选出与该用户记录编码特征比较相似的句子表征特征；再将这些被挑选出的句子表征特征以及该用户记录编码特征进行语义交互处理，得到至少一个内容交互特征，以使这些内容交互特征能够表示出这些被挑选出的句子表征特征以及该用户记录编码特征之间的语义交互结果，以便后续能够从这些被挑选出的句子表征特征所表征的待使用句子中，确定出与待使用用户记录语义最接近的句子，如此能够有效地减少语义交互对象个数，从而有利于提高语义交互效率，进而有利于提高纪要确定效率。

基于上述S322的第三种可能的实施方式的相关内容可知，在获取到用户记录编码特征与至少一个句子编码特征之后，可以先参考各个待使用句子的上下文信息，对各个句子编码特征进行句编码处理，得到各个句子表征特征；再针对用户记录编码特征、以及全部或者部分句子表征特征进行语义交互处理，得到至少一个内容交互特征，以便后续能够基于这些内容交互特征，更好地确定出与待使用用户记录语义最接近的句子，如此有利于提高纪要确定效果。

另外，本申请实施例不限定上述“内容交互模块”的实施方式，例如，其具体可以包括：注意力层和融合层。

为了便于理解上述“内容交互模块”的工作原理，下文以上述“至少一个内容交互特征”的确定过程为例进行说明。

作为示例，当上述“内容交互模块”的输入数据包括用户记录编码特征以及至少一个待交互句子特征时，利用上述“内容交互模块”确定“至少一个内容交互特征”的过程包括步骤31-步骤32：

步骤31：将用户记录编码特征以及至少一个待交互句子特征输入注意力层，得到该注意力层输出的用户记录表征特征。

上述“至少一个待交互句子特征”用于代表在对用户记录编码特征进行语义交互处理时所需参考的语句特征；而且本申请实施例不限定该“至少一个待交互句子特征”，例如，当具有步骤31-步骤32所示的工作原理的“内容交互模块”被用于实现上文“S322的第一种可能的实施方式”时，该“至少一个待交互句子特征”可以包括上文“至少一个句子编码特征”。又如，当具有步骤31-步骤32所示的工作原理的“内容交互模块”被用于实现上文步骤12时，该“至少一个待交互句子特征”包括上文“至少一个待处理句子特征”。还如，当具有步骤31-步骤32所示的工作原理的“内容交互模块”被用于实现上文“步骤22的第一种可能的实施方式”时，该“至少一个待交互句子特征”包括上文“至少一个句子表征特征”。也如，当具有步骤31-步骤32所示的工作原理的“内容交互模块”被用于实现上文步骤222时，该“至少一个待交互句子特征”包括上文“至少一个待参考句子特征”。

上述“注意力层”用于针对该注意力层的输入数据进行注意力处理；而且本申请实施例不限定上述“注意力层”的实施方式，例如，可以采用多头注意力机制进行实施。

上述“用户记录表征特征”用于表示待使用用户记录携带的语义信息。

另外，本申请实施例不限定上述“用户记录表征特征”的确定过程，例如，当上述“注意力层”采用多头注意力机制进行实施时，该“用户记录表征特征”的确定过程具体可以包括：将用户记录编码特征作为该多头注意力机制的询问参数（也就是，参数Q），并将上述“至少一个待交互句子特征”作为该多头注意力机制的其他两个参数（也就是，参数K和参数V），如此能够将述“至少一个待交互句子特征”携带的会议内容信息融入到用户记录编码特征中，以使在待使用用户记录较为简短，或者描述模糊等信息不全的情况下，这些会议内容信息能够较好地补充出该待使用用户记录中省略的信息，同时还使得该“用户记录表征特征”拥有了全场会议内容的全局视野，如此有利于更好地确定出该待使用用户记录在上述“待使用记录文本”中所对应的语句。

步骤32：根据用户记录表征特征、至少一个待交互句子特征、以及融合层，确定至少

一个内容交互特征。

其中，“第 e 个内容交互特征”是指用户记录表征特征与第 e 个待交互句子特征之间的语义交互结果。其中，e 为正整数，e≤E，E 为正整数，E 表示内容交互特征的个数（也就是，待交互句子特征的个数）。

另外，本申请实施例不限定上述“第 e 个内容交互特征”的确定过程，例如，其具体可以包括：将用户记录表征特征与第 e 个待交互句子特征输入融合层，得到融合层输出的第 e 个内容交互特征。

上述“融合层”用于将该融合层的输入数据进行特征融合处理；而且本申请实施例不限定该“融合层”的工作原理，例如，其具体可以为：在将用户记录表征特征与第 e 个待交互句子特征输入融合层之后，该融合层可以直接将该用户记录表征特征与第 e 个待交互句子特征进行拼接处理，得到第 e 个内容交互特征。

另外，为了提高语义交互效果，本申请实施例还提供了上述“融合层”的另一种工作原理，为了便于理解，下面以上述“第 e 个内容交互特征”的确定过程为例进行说明。

作为示例，“第 e 个内容交互特征”的确定过程，具体可以包括步骤 321-步骤 323：

步骤 321：根据用户记录表征特征和第 e 个待交互句子特征，确定第 e 个相似度表征特征，以使该第 e 个相似度表征特征用于表示该用户记录表征特征与该第 e 个待交互句子特征之间的信息相似性。

上述“第 e 个相似度表征特征”用于描述用户记录表征特征和第 e 个待交互句子特征之间的语义信息相似性。

另外，本申请实施例不限定上述“第 e 个相似度表征特征”的确定过程，例如，可以采用现有的或者未来出现的任一种特征相似性确定方法进行实施。又如，其具体可以为：根据用户记录表征特征与第 e 个待交互句子特征之间的点积，确定第 e 个相似度表征特征（例如，将用户记录表征特征与第 e 个待交互句子特征之间的点积，确定为第 e 个相似度表征特征）。

步骤 322：根据用户记录表征特征和第 e 个待交互句子特征，确定第 e 个信息差表征特征，以使该第 e 个信息差表征特征用于表示该用户记录表征特征与该第 e 个待交互句子特征之间的信息差异性。

上述“第 e 个信息差表征特征”用于描述用户记录表征特征和第 e 个待交互句子特征之间的语义信息差异性，以使该“第 e 个信息差表征特征”能够表示出待使用用户记录以及该第 e 个待交互句子特征所表征的待使用句子之间的信息差。

另外，本申请实施例不限定上述“第 e 个信息差表征特征”的确定过程，例如，可以采用现有的或者未来出现的任一种特征差异性确定方法进行实施。又如，其具体可以为：根据用户记录表征特征与第 e 个待交互句子特征之间的差值，确定第 e 个信息差表征特征（例如，将用户记录表征特征与第 e 个待交互句子特征之间的差值，确定为第 e 个信息差表征特征）。

步骤 323：将用户记录表征特征、第 e 个待交互句子特征、第 e 个相似度表征特征、以及第 e 个信息差表征特征进行拼接处理，得到第 e 个内容交互特征（如公式（1）所示）。

$$s_e'' = [q', \widetilde{s_e}, q' * \widetilde{s_e}, q' - \widetilde{s_e}] \quad (1)$$

式中，s_e''表示第 e 个内容交互特征；q'表示用户记录表征特征；$\widetilde{s_e}$表示第 e 个待交互句子特征；*表示点积处理；$q' * \widetilde{s_e}$表示第 e 个相似度表征特征；$q' - \widetilde{s_e}$表示第 e 个信息差表征特征。

基于上述步骤 321 至步骤 323 的相关内容可知，对于上述“融合层”来说，在获取到用户记录表征特征以及第 e 个待交互句子特征之后，可以参考该用户记录表征特征、第 e 个待交互句子特征、该用户记录表征特征与第 e 个待交互句子特征之间的相似性特征、以及该用户记录表征特征与第 e 个待交互句子特征之间的差异性特征，确定第 e 个内容交互特征，以使该第 e 个内容交互特征能够更全面地描述出待使用用户记录与该第 e 个待交互句子特征所表征的待使用句子之间的语义相似性。

基于上述步骤 31 至步骤 32 的相关内容可知，对于上述“内容交互模块”来说，其可以针对该内容交互模块的输入数据，先进行注意力处理再进行特征融合处理，以实现针对该内容交互模块的输入数据的语义交互处理。

需要说明的是，在利用步骤 31-步骤 32 实现上文“S322 的第一种可能的实施方式”时，可以直接将步骤 31-步骤 32 中“待交互句子特征”替换为“句子编码特征”即可；在利用步骤 31-步骤 32 实现上文步骤 12 时，可以直接将步骤 31-步骤 32 中“待交互句子特征”替换为“待处理句子特征”即可；在利用步骤 31-步骤 32 实现上文“步骤 22 的第一种可能的实施方式”时，可以直接将步骤 31-步骤 32 中“待交互句子特征”替换为“句子表征特征”即可；在利用步骤 31-步骤 32 实现上文步骤 222 时，可以直接将步骤 31-步骤 32 中“待交互句子特征”替换为“待参考句子特征”即可。

基于上述 S322 的相关内容可知，对于语义匹配模型来说，在利用该语义匹配模型中编码模块，确定出用户记录编码特征以及至少一个句子编码特征之后，可以由该语义匹配模型中内容交互模块参考该用户记录编码特征以及这些句子编码特征进行语义交互处理，得到至少一个内容交互特征，以使这些内容交互特征能够更好地表示出待使用用户记录与全部或者部分待使用句子之间的语义相似性。

S323：将至少一个内容交互特征输入匹配预测模块，得到该匹配预测模块输出的匹配预测结果。

上述“匹配预测模块”用于确定某个待使用句子携带的语义信息与待使用用户记录携带的语义信息之间的匹配程度（也就是，相似程度）；而且本申请实施例不限定该“匹配预测模块”的实施方式。

上述“匹配预测结果”用于表示上述“至少一个待使用句子”与待使用用户记录之间的语义匹配结果；而且本申请实施例不限定该“匹配预测结果”，例如，若第 e 个内容交互特征是根据用户记录表征特征以及第 e 个待交互句子特征进行确定的，则该“匹配预测结果”可以包括：待使用用户记录与该第 e 个待交互句子特征所表征的待使用句子之间的匹配预测得分，以使该“待使用用户记录与该第 e 个待交互句子特征对应的待使用句子之间的匹配预测得分”能够表示出该待使用用户记录与该第 e 个待交互句子特征所表征的待使用句子之间的语义相似程度。其中，e 为正整数，e≤E，E 为正整数。

基于上述 S321 至 S323 的相关内容可知，对于预先构建的语义匹配模型来说，其可以借助该语义匹配模型中词编码模块、内容交互模块以及匹配预测模块，实现针对待使用用

户记录以及待使用匹配文本中至少一个语句之间的语义匹配过程。其中，因上述“语义匹配模型”具有较好的语义匹配效果，使得基于该“语义匹配模型”确定的匹配预测结果能够更好地表示出表示上述“至少一个待使用句子”与待使用用户记录之间的语义相似程度。

另外，上述“语义匹配模型”可以根据预先收集的记录语义匹配标注语料进行构建；而且本申请实施例不限定上述“语义匹配模型”的构建过程，例如，其可以采用现有的或者未来出现的任意一种模型构建方法进行实施。

上述“记录语义匹配标注语料”可以包括由样本用户记录、该样本用户记录对应的第一会议文本句子集合、该样本用户记录对应的语义匹配标签。其中，“样本用户记录”是指由相关人员针对样本会议记录的速写信息。“样本用户记录对应的第一会议文本句子集合”是通过针对样本会议的会议转写文本进行句子切分得到的。“样本用户记录对应的语义匹配标签”用于表示样本用户记录的语义信息实际上与该第一会议文本句子集合中哪些句子的语义信息相同，而且该“样本用户记录对应的语义匹配标签”可以由人工标注的方式生成。

此外，为了提高上述“语义匹配模型”的抗噪性能，可以通过针对上述“记录语义匹配标注语料”进行数据加噪处理，构造一些弱监督数据，以便后续能够利用这些弱监督数据以及上述“记录语义匹配标注语料”，对上述“语义匹配模型”进行构建。

上述“弱监督数据”的构造过程具体可以包括：先对于上述“样本会议的会议转写文本”中每个原始句子，通过同义词替换、词片段删除，随机词片段插入，以及词片段乱序等操作，生成各个原始句子的句子加噪数据；再利用该“样本会议的会议转写文本”中各个原始句子与各个原始句子的句子加噪数据构成一对具有相关性的句对，以便后续能够基于这些句对，对上述“语义匹配模型”进行构建；而且该“语义匹配模型”的构架过程包括：先利用这些句对该语义匹配模型进行第一阶段训练，以使训练好的语义匹配模型能够准确地学习到该句对中句子加噪数据携带的语义信息与该句对中原始句子携带的语义信息相同；再利用上述“记录语义匹配标注语料”对语义匹配模型进行第二阶段训练，以使该语义匹配模型针对样本用户记录以及第一会议文本句子集合输出的语义预测结果几乎接近于上述“样本用户记录对应的语义匹配标签”，如此有利于提高语义匹配模型的语义匹配性能。

基于上述 S32 的相关内容可知，在获取到待使用用户记录以及待使用匹配文本之后，可以利用预先构建的语义匹配模型，针对该待使用用户记录以及该待使用匹配文本中至少一个语句进行语义匹配处理，得到匹配预测结果，以使该“匹配预测结果”能够表示上述“至少一个待使用句子”与待使用用户记录之间的语义相似程度，以便后续能够基于该“匹配预测结果”，从待使用记录文本中确定出与该待使用用户记录语义相似的语句。

S33：根据匹配预测结果，确定待使用语义匹配结果。

本申请实施例不限定 S33 的实施方式，例如，其具体可以包括：将匹配预测结果，确定为待使用语义匹配结果。

另外，为了进一步提高语义匹配准确性，本申请实施例还提供了 S33 的另一种可能的实施方式，其具体可以包括：根据待使用文本检索结果以及匹配预测结果，确定待使用语义匹配结果。

上述“待使用文本检索结果”是指从上述“待使用记录文本”中对待使用用户记录进

行检索处理得到的结果；而且本申请实施例不限定该“待使用文本检索结果”，例如，其可以包括：待使用用户记录与第 j 个待使用句子之间的检索得分；而且该“待使用用户记录与第 j 个待使用句子之间的检索得分”能够表示出该待使用用户记录与第 j 个待使用句子之间的语义相似程度。其中，j 为正整数，j≤J，J 为正整数。

另外，上述“待使用文本检索结果”可以根据预设检索算法、待使用用户记录以及至少一个待使用句子进行确定。其中，“预设检索算法”可以预先设定，例如，其可以是最佳匹配（Best Match，BM）25。需要说明的是，上述“25”指的是第 25 次算法迭代。

为了便于理解上述“S33 的另一种可能的实施方式”，下面结合示例进行说明。

作为示例，当上述“匹配预测结果”包括待使用用户记录与第 j 个待使用句子之间的匹配预测得分，且上述“待使用文本检索结果”包括该待使用用户记录与第 j 个待使用句子之间的检索得分（例如，BM25 检索得分）时，S33 具体可以为：将该“待使用用户记录与第 j 个待使用句子之间的匹配预测得分”、以及该“待使用用户记录与第 j 个待使用句子之间的检索得分”进行加权求和，得到该待使用用户记录与第 j 个待使用句子之间的语义匹配得分。其中，j 为正整数，j≤J，J 为正整数。

需要说明的是，上述“待使用用户记录与第 j 个待使用句子之间的匹配预测得分”对应的加权权重、以及上述“待使用用户记录与第 j 个待使用句子之间的检索得分”对应的加权权重均可以预先设定。

基于上述 S31 至 S33 的相关内容可知，在获取到待使用用户记录以及至少一个待使用句子之后，可以利用预先构建的语义匹配模型，实现针对该待使用用户记录以及至少一个待使用句子之间的语义匹配过程，得到待使用语义匹配结果。其中，因上述“语义匹配模型”具有较好的语义匹配效果，使得基于该“语义匹配模型”确定的待使用语义匹配结果，能够更好地表示出待使用用户记录与待使用记录文本中至少一个句子之间的语义匹配程度，如此有利于提高纪要确定效果。

方法实施例三

发明人还发现，一条用户记录信息在待使用记录文本中所对应的多个句子的位置分布通常比较集中，故基于该位置分布约束，本申请实施例还提供了纪要确定方法的另一种可能的实施方式，在该实施方式中，当上文“至少一个待使用句子”包括至少一个语义匹配句子，而且上文“待使用语义匹配结果”包括待使用用户记录与各语义匹配句子之间的语义匹配得分时，该纪要确定方法除了包括上述 S1-S4 以外，该纪要确定方法还包括 S5-S7，且 S5-S7 的执行顺序介于 S3 与 S4 之间。

S5：根据待使用用户记录与各语义匹配句子之间的语义匹配得分，从至少一个语义匹配句子中筛选出满足预设基准条件的至少一个基准句子。

上述“语义匹配句子”是指具有语义匹配得分的待使用句子（也就是，上文“内容交互特征”所表征的待使用句子）；而且本申请实施例不限定该“语义匹配句子”，为了便于理解，下面结合三个示例进行说明。

示例 1，若在上文“待使用语义匹配结果”的确定过程中不存在筛选处理（例如，基于上文“第一相似条件”或者“第二相似条件”的筛选过程），则该“语义匹配句子”是指

上文“待使用句子”，以使上述“至少一个语义匹配句子”可以包括上文“至少一个待使用句子”。

示例 2，若在上文“待使用语义匹配结果”的确定过程中存在基于第一相似条件的筛选过程，则该“语义匹配句子”就是指待参考句子特征所表征的待使用句子，以使上述“至少一个语义匹配句子”可以包括至少一个待参考句子特征所表征的待使用句子。

示例 3，若在上文“待使用语义匹配结果”的确定过程中存在基于第二相似条件的筛选过程，则该“语义匹配句子”就是指待处理句子特征所表征的待使用句子，以使上述“至少一个语义匹配句子”可以包括至少一个待处理句子特征所表征的待使用句子。

上述“预设基准条件”可以预先设定，例如，其具体可以为：待使用用户记录与基准句子之间的语义匹配得分高于预先设定的第一得分阈值。又如，当将所有语义匹配句子按照语义匹配得分从大到小进行排序时，该“预设基准条件”具体可以为：排列序号低于预先设定的第三序号阈值（例如，2）。

上述“基准句子”是指满足预设基准条件的语义匹配句子。

基于上述 S5 的相关内容可知，在获取到待使用用户记录与各语义匹配句子之间的语义匹配得分之后，可以基于这些语义匹配得分，从所有语义匹配句子中筛选出语义匹配得分较高的多个语义匹配句子（例如，Top2），以便后续能够以这些语义匹配得分较高的多个语义匹配句子的文本位置作为基准，确定各个语义匹配句子的文本位置惩罚项。

S6：根据至少一个基准句子的文本位置、以及各语义匹配句子的文本位置，确定各语义匹配句子的距离惩罚得分。

其中，“第 e 个语义匹配句子的距离惩罚得分”用于表示该第 e 个语义匹配句子的文本位置对“待使用用户记录与第 e 个语义匹配句子之间的语义匹配程度”所造成的影响。其中，e 为正整数，e≤E。

另外，本申请实施例不限定上述“第 e 个语义匹配句子的距离惩罚得分”的确定过程，例如，其具体可以包括 S61-S63：

S61：根据各个基准句子的文本位置与第 e 个语义匹配句子的文本位置之间的距离，确定该第 e 个语义匹配句子对应的距离集。

例如，当上述“基准句子”的个数为 Y 时，S61 具体可以包括：将第 1 个基准句子的文本位置与第 e 个语义匹配句子的文本位置之间的距离、第 2 个基准句子的文本位置与第 e 个语义匹配句子的文本位置之间的距离、……、以及第 Y 个基准句子的文本位置与第 e 个语义匹配句子的文本位置之间的距离进行集合处理，得到该第 e 个语义匹配句子对应的距离集。

需要说明的是，上述“文本位置”是指一个句子在待使用记录文本中所处位置。例如，上述“第 e 个语义匹配句子的文本位置”是指该第 e 个语义匹配句子在待使用记录文本中所处位置。

S62：对第 e 个语义匹配句子对应的距离集进行预设统计分析处理，得到该第 e 个语义匹配句子对应的待使用距离。

上述“预设统计分析处理”可以预先设定；而且本申请实施例不限定该“预设统计分析处理”，例如，其可以为取最小值、取最大值、取平均值、取和值等。

上述“第 e 个语义匹配句子对应的待使用距离”用于表示该第 e 个语义匹配句子与上述“至少一个基准句子”之间的文本位置距离。

S63：根据第 e 个语义匹配句子对应的待使用距离，确定该第 e 个语义匹配句子的距离惩罚得分。

需要说明的是，本申请实施例不限定 S63 的实施方式，例如，其具体可以包括：将预设惩罚系数与第 e 个语义匹配句子对应的待使用距离之间的乘积，确定为该第 e 个语义匹配句子的距离惩罚得分。其中，“预设惩罚系数”可以预先设定。

基于上述 S6 的相关内容可知，在获取到至少一个基准句子之后，可以根据各个语义匹配句子的文本位置与这些基准句子的文本位置之间的距离，确定各个语义匹配句子的距离惩罚得分。

S7：利用各语义匹配句子的距离惩罚得分，更新待使用用户记录与各语义匹配句子之间的语义匹配得分（如公式（2）所示）。

$$Score'_e = Score_e - DP_e \quad (2)$$

式中，$Score'_e$表示更新后的待使用用户记录与第 e 语义匹配句子之间的语义匹配得分；$Score_e$表示更新前的待使用用户记录与第 e 语义匹配句子之间的语义匹配得分；DP_e表示第 e 个语义匹配句子的距离惩罚得分。

基于上述 S5 至 S7 的相关内容可知，在获取到待使用语义匹配结果之后，可以利用该待使用语义匹配结果所涉及的各个待使用句子的文本位置，对该待使用语义匹配结果进行更新处理，以使更新后的待使用语义匹配结果满足实际用户记录描述语句较为集中的位置分布约束，从而使得更新后的待使用语义匹配结果能够更准确地表示出待使用用户记录与待使用记录文本中至少一个句子之间的语义匹配程度，以便后续能够基于更新后的待使用语义匹配结果，确定待使用纪要内容，如此有利于提高纪要确定效果。

方法实施例四

另外，为了进一步提高纪要确定效果，本申请实施例还提供了上述 S4 的另一种可能的实施方式，下面结合示例进行说明。

作为示例，当上文“至少一个待使用句子”包括至少一个语义匹配句子，而且上文“待使用语义匹配结果”包括待使用用户记录与各语义匹配句子之间的语义匹配得分时，S4 具体可以包括 S41-S42：

S41：根据待使用用户记录与各语义匹配句子之间的语义匹配得分，从至少一个语义匹配句子中筛选出满足第一匹配条件的至少一个目标句子。

上述“第一匹配条件”可以预先设定，例如，其可以采用下文 S41 的任一实施方式所示的“至少一个目标句子”所达到的条件（例如，下文公式（6）所示的条件）进行实施。

上述“目标句子”是指满足第一匹配条件的语义匹配句子。

本申请实施例不限定 S41 的实施方式，例如，其可以采用下文步骤 41-步骤 42 所示的“至少一个目标句子”的确定过程的任一实施方式进行实施，只需将下文步骤 41-步骤 42 所示的“至少一个目标句子”的确定过程的任一实施方式中“候选句子”替换为“语义匹配句子”即可。

另外，本申请实施例还提供了 S41 的另一种可能的实施方式，其具体可以包括 S411-S412：

S411：根据待使用用户记录与各语义匹配句子之间的语义匹配得分，从至少一个语义匹配句子中筛选出满足第二匹配条件的至少一个初选句子。

上述“第二匹配条件”可以预先设定，例如，其具体可以为：待使用用户记录与初选句子之间的语义匹配得分高于预先设定的第二得分阈值。又如，当将所有语义匹配句子按照语义匹配得分从大到小进行排序时，该“第二匹配条件”具体可以为：排列序号低于预先设定的第四序号阈值。

上述“初选句子”是指满足第二匹配条件的语义匹配句子。

基于 S411 的相关内容可知，在获取到待使用用户记录与各语义匹配句子之间的语义匹配得分之后，可以基于这些语义匹配得分，从所有语义匹配句子中筛选出语义匹配得分较高的多个语义匹配句子（例如，TopK），分别确定为初选句子，以使这些初选句子携带的语义信息比较接近于待使用用户记录携带的语义信息，以便后续能够从这些初选句子中确定出用于表示该待使用用户记录所携带的语义信息的句子。

S412：根据待使用用户记录与各初选句子之间的语义匹配得分，从至少一个初选句子中筛选出满足第三匹配条件的至少一个目标句子。

上述“第三匹配条件”可以预先设定，例如，其具体可以为：下文步骤 41-步骤 42 所示的“至少一个目标句子”所达到的条件（例如，下文公式（6）所示的条件）。

本申请实施例不限定 S412 的实施方式，例如，其可以采用下文步骤 41-步骤 42 所示的“至少一个目标句子”的确定过程的任一实施方式进行实施，只需将下文步骤 41-步骤 42 所示的“至少一个目标句子”的确定过程的任一实施方式中“候选句子”替换为“初选句子”即可。

基于上述 S411 至 S412 的相关内容可知，在获取到待使用语义匹配结果之后，可以先借助该待使用语义匹配结果所记录的语义匹配得分，确定出满足第二匹配条件的各个初选句子；再将满足第三匹配条件的各个初选句子，分别确定为目标句子，以使该目标句子能够同时满足第二匹配条件以及第三匹配条件，从而使得这些目标句子能够更好地表示出待使用用户记录所携带的语义信息。

此外，本申请实施例还提供了 S41 的又一种可能的实施方式，在该实施方式中，S41 不仅包括上述 S411，可以还包括 S413-S414：

S413：根据各初选句子的非停用词个数，从至少一个初选句子中筛选出满足第四匹配条件的至少一个再选句子。

上述“非停用词个数”是指一个句子中除了停用词以外其他所有词汇的个数。

上述“第四匹配条件”可以预先设定，例如，其具体可以为：非停用词个数超过预设个数阈值（例如，1）。

上述“再选句子”是指满足第四匹配条件的初选句子（也就是，既满足第二匹配条件又满足第四匹配条件的语义匹配句子）。

基于 S413 的相关内容可知，在获取到至少一个初选句子之后，可以先统计各个初选句子的非停用词个数；再将非停用词个数比较高的各个初选句子，分别确定为再选句子，如

此能够实现针对低信息量句子的过滤处理，从而能够有效地避免这些低信息量句子对纪要确定过程所造成的不良影响。

S414：根据待使用用户记录与各再选句子之间的语义匹配得分，从至少一个再选句子中筛选出满足第三匹配条件的至少一个目标句子。

需要说明的是，本申请实施例不限定 S414 的实施方式，例如，其可以采用下文步骤 41-步骤 42 所示的“至少一个目标句子”的确定过程的任一实施方式进行实施，只需将下文步骤 41-步骤 42 所示的“至少一个目标句子”的确定过程的任一实施方式中“候选句子”替换为“再选句子”即可。

基于上述 S413 至 S414 的相关内容可知，在获取到待使用语义匹配结果之后，可以先借助该待使用语义匹配结果所记录的语义匹配得分，确定出满足第二匹配条件的各个初选句子；再从这些初选句子中滤除低信息量句子，得到至少一个再选句子；最后，将满足第三匹配条件的各个再选句子，分别确定为目标句子，以使该目标句子能够同时满足第二匹配条件、第三匹配条件以及第四匹配条件，从而使得这些目标句子能够更好地表示出待使用用户记录所携带的语义信息。

另外，为了能够更好地确定出上述“至少一个目标句子”，本申请实施例还提供了确定上述“至少一个目标句子”的一种可能的实施方式，为了便于理解，下面结合示例进行说明。

作为示例，当上述“至少一个目标句子”是从至少一个候选句子中筛选出来的时，该“至少一个目标句子”的确定过程，具体可以包括步骤 41-步骤 42：

步骤 41：根据待使用用户记录与各候选句子之间的语义匹配得分，从至少一个候选句子中筛选出满足第一丢弃条件的待删除句子，得到第一筛选结果。

上述“至少一个候选句子”包括上述“至少一个目标句子”；而且本申请实施例不限定上述“至少一个候选句子”，例如，当上文 S41 采用步骤 41-步骤 42 进行实施时，上述“至少一个候选句子”是指上文“至少一个语义匹配句子”。又如，当上文 S412 采用采用步骤 41-步骤 42 进行实施时，上述“至少一个候选句子”是指上文“至少一个初选句子”。还如，当上文 S414 采用采用步骤 41-步骤 42 进行实施时，上述“至少一个候选句子”是指上文“至少一个再选句子”。

上述“第一丢弃条件”可以预先设定，例如，其具体可以为：待使用用户记录与待删除句子之间的语义匹配得分低于预先设定的第三得分阈值（例如，0.6）。

上述“待删除句子”是指满足第一丢弃条件的候选句子。

上述“第一筛选结果”用于表示上述“至少一个候选句子”中是否存在待删除句子。

基于上述步骤 41 的相关内容可知，在获取到至少一个候选句子之后，可以从这些候选句子中挑选出语义匹配得分比较低的候选句子，分别确定为待删除句子，以便后续能够从这些待删除句子中确定出那些需要删除的句子。

步骤 42：根据第一筛选结果和至少一个候选句子，确定至少一个目标句子。

作为示例，步骤 42 具体可以包括步骤 421-步骤 424：

步骤 421：根据第一筛选结果，确定上述“至少一个候选句子”中是否存在待删除句子，若是，则执行步骤 423-步骤 424；若否，则执行步骤 422。

本申请实施例中，在获取到第一筛选结果之后，若该第一筛选结果表示上述“至少一个候选句子”中不存在待删除句子，故可以确定所有候选句子的语义匹配得分均比较高，从而可以确定所有候选句子携带的语义信息十分接近于待使用用户记录携带的语义信息；若该第一筛选结果表示上述“至少一个候选句子”中存在至少一个待删除句子，故可以确定上述“至少一个候选句子”中存在一些语义匹配得分比较低的候选句子，从而可以确定这些语义匹配得分比较低的候选句子所携带的语义信息与待使用用户记录携带的语义信息之间差异性可能比较大。

步骤422：可以将各个候选句子均确定为目标句子。

本申请实施例中，若第一筛选结果表示上述“至少一个候选句子”中不存在待删除句子，则可以确定所有候选句子携带的语义信息十分接近于待使用用户记录携带的语义信息，故可以直接将各个候选句子均确定为目标句子，以使这些目标句子能够表示出该待使用用户记录在待使用记录文本中所对应的句子。

步骤423：根据至少一个待删除句子和至少一个候选句子，确定至少一个待保留句子。

上述“待保留句子”是指需要保留下来的句子（也就是，无需删除的句子）。

本申请实施例不限定步骤423的实施方式，例如，其具体可以为：若第一筛选结果表示上述“至少一个候选句子”中存在至少一个待删除句子，则可以从该“至少一个候选句子”中删除所有待删除句子，得到至少一个待保留句子。

在一些情况下，因待使用用户记录存在缺陷（例如，过于简略或者概括性较强），使得基于该待使用用户记录筛选所得的所有候选句子的语义匹配得分均比较低，故为了提高用户记录信息的定位准确性，本申请实施例还提供了步骤423的另一种可能的实施方式，其具体可以包括步骤4231-步骤4235：

步骤4231：若至少一个候选句子满足预设校正条件，则从至少一个待删除句子中筛选出满足第一保留条件的至少一个待评估句子。

上述“预设校正条件”可以预先设定，例如，其具体可以为：待使用用户记录与各候选句子之间的语义匹配得分均不高于预先设定的第四得分阈值（也就是，待使用用户记录与各候选句子之间的语义匹配得分都比较低）。

上述“第一保留条件”可以预先设定，例如，其具体可以为：待使用用户记录与待评估句子之间的语义匹配得分高于预先设定的第五得分阈值（例如，0.3），而且该第五得分阈值低于上文“第三得分阈值”。

上述“待评估句子”是指满足第一保留条件的待删除句子。例如，当上文“第三得分阈值”为0.6，且上述“第五得分阈值”为0.3时，待使用用户记录与待评估句子之间的语义匹配得分介于0.6到0.3之间。

基于上述步骤4231的相关内容可知，在获取到至少一个待删除句子之后，若确定上述“至少一个候选句子”满足预设校正条件，则可以确定待使用用户记录存在缺陷，故可以将一些语义匹配得分比较高的待删除句子，分别确定为待评估句子，以便后续能够从这些待评估句子中挑选出一些需要保留的句子。

步骤4232：对待使用用户记录与各待评估句子之间的语义匹配得分进行归一化处理，得到各待评估句子的归一化得分。

需要说明的是，本申请实施例不限定步骤4232的实施方式，例如，可以采用现有的或者未来出现的任意一种数据归一化方法进行实施。又如，步骤4232可以采用公式（3）-（5）所示的归一化方法进行实施

$$\mu = \frac{\sum_{d=1}^{D} Score^{d}}{D} \tag{3}$$

$$\sigma = \sqrt{\frac{\sum_{d=1}^{D} (Score^{d} - \mu)^{2}}{D}} \tag{4}$$

$$score_norm_{l} = \frac{score^{l} - \mu}{\sigma} \tag{5}$$

式中，$score_norm_l$表示第l个待评估句子的归一化得分；$score^l$表示待使用用户记录与第l个待评估句子之间的语义匹配得分；l为正整数，l≤L，L为正整数，L表示待评估句子的个数；$Score^d$表示待使用用户记录与第d个候选句子之间的语义匹配得分；d为正整数，d≤D，D为正整数，D表示候选句子的个数；σ表示上述“至少一个候选句子”的语义匹配得分的方差值；μ表示上述“至少一个候选句子”的语义匹配得分的平均值。

步骤4233：根据各待评估句子的归一化得分，从至少一个待评估句子中筛选出满足第二保留条件的至少一个第一保留句子。

上述“第二保留条件”可以预先设定，例如，其具体可以包括：待使用用户记录与第一保留句子之间的语义匹配得分高于预先设定的第六得分阈值（例如，0.3）。

上述“第一保留句子”是指满足第二保留条件的待评估句子。

基于上述步骤4233的相关内容可知，在获取到各个待评估句子的归一化得分之后，可以将一些归一化得分比较高的待评估句子均确定为第一保留句子，以使这些第一保留句子携带的语义信息比较接近于待使用用户记录所携带的语义信息。

步骤4234：从至少一个候选句子中删除至少一个待删除句子，得到至少一个第二保留句子。

需要说明的是，本申请实施例不限定步骤4234与上文步骤4231之间的执行顺序，例如，可以依次执行步骤4234和上文步骤4231，也可以依次执行步骤4231与上文步骤4234，还可以同时执行步骤4234与上文步骤4231。

步骤4235：根据至少一个第一保留句子、以及至少一个第二保留句子，确定至少一个待保留句子。

本申请实施例中，在获取到至少一个第一保留句子以及至少一个第二保留句子之后，可以将各个第一保留句子以及各个第二保留句子均确定为待保留句子，以使这些待保留句子均满足公式（6）所示的条件，以便后续能够从这些待保留句子中挑选出待使用用户记录对应的目标句子。

$$(sc_b \geq t_1) \text{ or } ((t_1 > sc_b > t_2) \,\&\&\, (sc_norm_b > t_3)) \tag{6}$$

式中，sc_b表示待使用用户记录与第b个待保留句子之间的语义匹配得分；b为正整数，b≤B，B为正整数，B表示待保留句子的个数；sc_norm_b表示第b个待保留句子的归一化得分；t_1表示第三得分阈值；t_2表示第五得分阈值；t_3表示第六得分阈值。

基于上述步骤423的相关内容可知，若第一筛选结果表示上述“至少一个候选句子”中存在至少一个待删除句子，则可以参考这些待删除句子，从这些候选句子中挑选出一些待

保留句子，以使这些待保留句子携带的语义信息更接近于待使用用户记录所携带的语义信息。

步骤 424：根据至少一个待保留句子，确定至少一个目标句子。

需要说明的是，本申请实施例不限定步骤 424 的实施方式，例如，其具体可以为：将各个待保留句子均确定为目标句子。

基于上述步骤 41 至步骤 42 的相关内容可知，在获取到至少一个候选句子之后，可以按照预设句子挑选策略（也就是，步骤 41-步骤 42 所示的句子挑选策略），从这些候选句子中挑选出一个或者多个目标句子，以使这些目标句子能够表示出待使用用户记录所携带的语义信息。

基于上述 S41 的相关内容可知，在获取到待使用语义匹配结果之后，可以先参考该待使用语义匹配结果所记录的语义匹配得分，筛选出一些语义匹配得分排名比较靠前的语义匹配句子，分别确定为各个初选句子；再从这些初选句子中剔除那些非停用词个数小于等于 1 的低信息量句子，以得到一些信息量比较高的再选句子；然后，在确定这些再选句子对应的语义匹配得分均比较低时，可以按照公式（6）所示的条件，从这些再选句子中挑选出一个或者多个目标句子，以使这些目标句子能够表示出待使用用户记录所携带的语义信息。

S42：根据待使用用户记录和至少一个目标句子，确定待使用纪要内容。

需要说明的是，本申请实施例不限定 S42 的实施方式，例如，如图 2 所示，S42 具体可以为：可以将这些目标句子在待使用记录文本中突出显示给用户，并在这些目标句子的周围显示待使用用户记录，以使该用户能够基于这些显示内容编写待使用纪要内容，从而使得该用户能够从这些显示内容中获知其编写纪要所需内容，如此能够有效地避免因用户通过重听会议录制音频确定与该待使用用户记录相关的会议内容而导致的时长消耗，从而能够有效地提高纪要确定效率。

还需要说明的是，本申请实施例不限定上述“至少一个目标句子”的显示触发条件，例如，可以响应于用户触发的用户记录定位显示请求进行显示。其中，“用户记录定位显示请求”用于请求显示用户记录对应的目标句子在待使用记录文本中所处位置；而且本申请实施例不限定该“用户记录定位显示请求”的触发方式，例如，用户可以通过点击图 2 所示的望远镜按钮进行触发该“用户记录定位显示请求”。

另外，为了进一步提高纪要确定效果，S42 可以采用下文**方法实施例五**以及**方法实施例六**所示的 S42 的任一实施方式进行实施。

基于上述 S41 至 S42 的相关内容可知，在获取到待使用语义匹配结果之后，可以先基于待使用语义匹配结果，从待使用记录文本中确定出与待使用用户记录对应的目标句子；再根据这些目标句子以及待使用用户记录，确定待使用纪要内容。其中，因这些目标句子能够准确地表示出待使用用户记录所对应的会议记录内容，以使这些目标句子能够更准确地表示出该待使用用户记录携带语义信息，从而使得基于这些目标句子确定的待使用纪要内容更准确，如此有利于提高纪要确定效果。

方法实施例五

发明人还发现，在听取会议的过程中，相关人员需要一心多用，实时判断会议内容是否重要，并组织语言形成会议笔记。然而，因相关人员精力过于分散而导致会议笔记编写判断过程发生失误（例如，误判一些事项的重要程度），或者，因相关人员参与会议讨论或者暂时离开会议而导致记录遗漏，从而导致后续基于会议笔记所确定的会议纪要不准确，如此为最终形成高质量会议纪要带来困难。

基于上述发现，为了进一步提高纪要确定效果，本申请实施例还提供了上述 S42 的另一种可能的实施方式，其具体可以包括 S421-S422：

S421：根据至少一个目标句子、至少一个待使用句子、以及预先构建的补充语句确定模型，确定至少一个关键补充语句。

上述“关键补充语句”用于表示待使用记录文本中存在的待使用用户记录中未记录的比较重要的语义信息；而且上述“至少一个关键补充语句”可以用于表示该待使用用户记录遗漏的重要语义信息。

上述“补充语句确定模型”用于针对该补充语句确定模型的输入数据进行关键补充语句确定处理；而且该“补充语句确定模型”是一种机器学习模型（例如，基于图注意力网络（Graph attention network，GAT）结构的机器学习模型等）。

本申请实施例不限定 S421 的实施方式，例如，当上述“补充语句确定模型”是一种于图注意力网络（Graph attention network，GAT）结构的机器学习模型时，S421 具体可以包括 S4211- S4213：

S4211：根据至少一个目标句子和至少一个待使用句子，构建待使用图。

上述“待使用图”用于表示待使用记录文本中不同句子之间的关联关系；而且本申请实施例不限定该“待使用图”的构建过程，例如，其具体可以包括步骤 51-步骤 54：

步骤 51：根据至少一个目标句子和至少一个待使用句子，确定至少一个图节点，以使该“至少一个图节点”包括至少一个记录语句节点和至少一个候选语句节点。

作为示例，步骤 51 具体可以包括步骤 511-步骤 514：

步骤 511:构建各个记录语句节点,以使各个记录语句节点分别能够代表各个目标句子。

需要说明的是，本申请实施例不限定步骤 511 的实施方式，例如，其具体可以包括：将第 g 个目标句子，确定为第 g 个记录语句节点。又如，步骤 511 具体可以包括：先对该目标句子进行语义特征提取处理，得到该第 g 个目标句子的语义特征；再将该第 g 个目标句子的语义特征，确定为第 g 个记录语句节点，以使该第 g 个记录语句节点能够表示出该第 g 个目标句子携带的语义信息，从而使得该第 g 个记录语句节点能够代表该第 g 个目标句子。其中，g 为正整数，g≤G，G 为正整数，G 表示目标句子的个数。

上述“语义特征”用于表示一个句子携带的语义信息；而且本申请实施例不限定该“语义特征”的确定过程，例如，可以借助预先构建的语义特征提取模型（例如，如图 5 所示的 BERT）进行实施。

步骤 512：从至少一个待使用句子中删除至少一个目标句子，得到至少一个非记录语句。

步骤 513：根据各个非记录语句的语义特征，确定各个候选语句节点，以使各个候选语句节点分别能够代表各个非记录语句（也就是，上述“至少一个待使用句子”中除了上

述“至少一个目标句子”以外的各个待使用句子)。

需要说明的是，本申请实施例不限定步骤 513 的实施方式，例如，其具体可以包括：将第 a 个非记录语句，确定为第 a 个候选语句节点。又如，步骤 513 具体可以包括：先对该第 a 个非记录语句进行语义特征提取处理，得到该第 a 个非记录语句的语义特征；再将该第 a 个非记录语句的语义特征，确定为第 a 个候选语句节点，以使该第 a 个候选语句节点能够表示出该第 a 个非记录语句携带的语义信息，从而使得该第 a 个候选语句节点能够代表该第 a 个非记录语句。其中，a 为正整数，a≤A，A 为正整数，A 表示非记录语句的个数。

步骤 514：将所有记录语句节点以及所有候选语句节点进行集合处理，得到至少一个图节点，以使该“至少一个图节点”包括至少一个记录语句节点和至少一个候选语句节点。

基于上述步骤 51 的相关内容可知，在获取到至少一个目标句子之后，可以利用这些目标句子以及上文“至少一个待使用句子”，确定待使用图中各个节点，以使该待使用图中各个节点能够代表各个待使用句子。

步骤 52：按照预设近邻描述参数，构建至少一个图节点对应的近邻表征边。

上述“预设近邻描述参数”用于表示一个图像节点的近邻范围；而且本申请实施例不限定该“预设近邻描述参数”，例如，其可以包括：上下文窗口尺寸。

上述“上下文窗口尺寸”用于描述一个句子的上下文范围；而且本申请实施例不限定该“上下文窗口尺寸”，例如，其具体可以为 10（也就是，将与一个句子的文本位置相邻的 10 个句子，均确定为该句子的近邻)。

上述“近邻表征边”用于表示具有近邻关系的两个图节点之间的连接边，以使该“近邻表征边”用于表示两个图节点所代表句子之间的近邻关系。

另外，本申请实施例不限定上述“近邻表征边”的确定过程，例如，其具体可以包括：对于一个图节点来说，可以先根据预设近邻描述参数，确定该图节点所代表句子对应的上下文窗口；再将该图节点所代表句子对应的上下文窗口中任意两个句子所对应的图节点进行相连，得到这些图节点之间的近邻表征边。

基于上述步骤 52 的相关内容可知，在获取到至少一个图节点之后，可以先根据预设近邻描述参数，确定具有近邻关系的图节点对；再将每个图节点对所包括的两个图节点之间进行相连，得到这些图节点对应的近邻表征边（例如，图 5 中①所示的连接边)，以使这些近邻表征边能够表示不同图节点之间的近邻关系。

步骤 53：将各候选语句节点分别与各记录语句节点进行相连，得到各候选语句节点对应的相关表征边。

上述“相关表征边”用于表示一个候选语句节点与一个记录语句节点之间的连接关系。

基于步骤 53 的相关内容可知，在获取到至少一个记录语句节点和至少一个候选语句节点之后，可以将各个候选语句节点分别与各个记录语句节点进行相连，但是不能将不同记录语句节点之间进行相连，得到这些候选语句节点对应的相关表征边（例如，图 5 中②所示连接边)。

步骤 54：根据至少一个图节点、至少一个图节点对应的近邻表征边、以及至少一个候选语句节点对应的相关表征边，确定待使用图。

本申请实施例中，在获取到至少一个图节点、该至少一个图节点对应的近邻表征边、以及至少一个候选语句节点对应的相关表征边之后，可以参考这些图节点及其对应的连接边，构建待使用图，以使该待使用图包括该至少一个图节点、该至少一个图节点对应的近邻表征边、以及至少一个候选语句节点对应的相关表征边。

可见，对于待使用图中的任意一个图节点来说，如果该图节点属于记录语句节点（也就是，该图节点用于代表一个目标句子），则该图节点只对应于一种连接边（也就是，近邻表征边）；但是，如果该图节点属于候选语句节点（也就是，该图节点用于代表一个非记录语句），则该图节点可以对应于两种连接边（也就是，近邻表征边以及相关表征边）。

基于上述 S4211 的相关内容可知，在获取到至少一个目标句子之后，可以利用这些目标句子以及上文“至少一个待使用句子”，构建待使用图，以使该待使用图中各个节点分别代表各个待使用句子，并使得该待使用图中各个连接边能够表示出不同节点之间的连接关系，以便后续能够基于该待使用图进行关键补充语句确定处理。

S4212：将待使用图输入补充语句确定模型，得到该补充语句确定模型输出的关键语句预测结果。

上述“关键语句预测结果”用于表示各个待使用句子属于关键补充语句的可能性（尤其是，各个非记录语句属于关键补充语句的可能性）。

本申请实施例不限定上述“关键语句预测结果”的确定过程，例如，当上述“补充语句确定模型”包括图注意力模块、以及分类预测层时，该“关键语句预测结果”的确定过程，具体可以包括步骤 61-步骤 63：

步骤 61：确定待使用图中各个图节点的待使用特征。

其中，“第 r 个图节点的待使用特征”用于表示该第 r 个图节点所代表的待使用句子携带的语义信息。其中，r 为正整数，r≤R，R 为正整数，R 表示图节点的个数。

另外，本申请实施例不限定上述“第 r 个图节点的待使用特征”的确定过程，为了便于理解，下面结合两种情况进行说明。

情况 1，当第 r 个图节点采用第 r 个待使用句子的语义特征进行表示时，该“第 r 个图节点的待使用特征”就是该第 r 个图节点（也就是，该第 r 个待使用句子的语义特征）。

情况 2，当第 r 个图节点采用第 r 个待使用句子进行表示时，上述“补充语句确定模型”可以还包括编码层，而且该“第 r 个图节点的待使用特征”的确定过程，具体可以为：利用该编码层，对该第 r 个图节点（也就是，第 r 个待使用句子）进行编码处理，得到该第 r 个图节点的待使用特征。

上述“编码层”用于针对该编码层的输入数据进行编码处理；而且本申请实施例不限定该“编码层”的实施方式，例如，其可以采用 BERT 预训练模型进行实施。

步骤 62：将至少一个图节点的待使用特征以及待使用图输入图注意力模块，得到该图注意力模块输出的至少一个图节点的图注意特征。

上述“图注意力模块”用于针对该图注意力模块的输入数据进行图注意力处理；而且本申请实施例不限定该“图注意力模块”的实施方式，例如，其可以包括 M 个图注意力层（例如，2 个图注意力层）。其中，第 1 个图注意力层的输入数据包括上述“至少一个图节点的待使用特征”以及待使用图；而且第 m_1 个图注意力层的输入数据包括第 m_1-1 个图注

意力层的输出数据以及待使用图，m_1为正整数，$2 \leqslant m_1 \leqslant M$，M为正整数。

其中，“第m_2个图注意力层”用于针对该第m_2个图注意力层的输入层进行图注意力处理；而且本申请实施例不限定该“第m_2个图注意力层”的工作原理，为了便于理解，下面结合示例进行说明。其中，m_2为正整数，$1 \leqslant m_2 \leqslant M$。

作为示例，当第m_2个图注意力层的输入数据包括R个图节点的处理前特征以及待使用图，而且该第m_2个图注意力层的输出数据包括R个图节点的处理后特征时，第r个图节点的处理后特征确定过程（也就是，上述“第m_2个图注意力层”的工作原理），具体可以包括步骤71-步骤74：

步骤71：根据第r个图节点对应的近邻表征边，确定该第r个图节点对应的至少一个近邻节点。其中，r为正整数，$r \leqslant R$。

本申请实施例中，对于第r个图节点来说，可以将与该第r个图节点相连的各个近邻表征边的另一个端点，分别确定为该第r个图节点对应的各个近邻节点。可见，第r个图节点与该第r个图节点对应的各个近邻节点之间均被近邻表征边连接。

步骤72：根据第r个图节点的处理前特征、以及该第r个图节点对应的第d_r个近邻节点的处理前特征，确定该第d_r个近邻节点的近邻注意力权重特征。其中，d_r为正整数，$d_r \leqslant D_r$，D_r为正整数，D_r表示第r个图节点对应的近邻节点个数。

上述“第r个图节点的处理前特征”是指上述“第m_2个图注意力层的输入数据”中所包括的用于描述该第r个图节点的特征表示。

另外，本申请实施例不限定上述“第r个图节点的处理前特征”，例如，若$m_2=1$，则该“第r个图节点的处理前特征”可以是该r个图节点的待使用特征；若$m_2 \geqslant 2$，则该“第r个图节点的处理前特征”可以是上述“第m_2-1个图注意力层的输出数据”中所包括的用于描述该第r个图节点的特征表示（也就是，针对该r个图节点的待使用特征进行m_2-1次图注意力处理后得到的特征表示）。

上述“第d_r个近邻节点的近邻注意力权重特征”是指由上述“第r个图节点的处理前特征”与上述“第r个图节点对应的第d_r个近邻节点的处理前特征”之间进行语义交互得到的；而且本申请实施例不限定该“第d_r个近邻节点的近邻注意力权重特征”的确定过程，例如，其可以采用公式（7）进行实施。

$$\alpha_{d_r}^{r}=\frac{\exp(ReLu(\theta_1^{T}[W_1h_r||W_1h_{d_r}]))}{\sum_{x=1}^{D_r}\exp(ReLu(\theta_1^{T}[W_1h_r||W_1h_x]))} \tag{7}$$

式中，$\alpha_{d_r}^{r}$表示第r个图节点对应的第d_r个近邻节点的近邻注意力权重特征；h_r表示第r个图节点的处理前特征；h_{d_r}表示第r个图节点对应的第d_r个近邻节点的处理前特征；h_x表示第r个图节点对应的第x近邻节点的处理前特征，且x为正整数，$x \leqslant D_r$；W_1、θ_1均属于第m_2个图注意力层的层参数，且这些层参数可以在上述“补充语句确定模型”的构建过程中确定；$ReLu(\,)$表示激活函数。

需要说明的是，本申请实施例不限定上述公式（7）的实现方式，例如，可以借助注意力层、激活函数层以及归一化处理层进行实现。

步骤73：根据第r个图节点对应的D_r个近邻节点的处理前特征、以及该第r个图节点对应的D_r个近邻节点的近邻注意力权重特征，确定该第r个图节点的近邻融合特征。

需要说明的是，本申请实施例不限定步骤73的实施方式，例如，其具体可以采用公式（8）进行实施。

$$\hat{h}_r = \delta\left(\sum_{d_r=1}^{D_r} \alpha_{d_r}^{r} W_2 h_{d_r}\right) \quad (8)$$

式中，$\hat{h}_r$表示第r个图节点的近邻融合特征；$\alpha_{d_r}^{r}$表示第r个图节点对应的第d_r个近邻节点的近邻注意力权重特征；h_{d_r}表示第r个图节点对应的第d_r个近邻节点的处理前特征；W_2属于第m_2个图注意力层的层参数；$\delta(\ )$表示sigmoid函数。

步骤74：根据第r个图节点的近邻融合特征，确定该第r个图节点的处理后特征。

本申请实施例不限定步骤74的实施方式，例如，其可以采用公式（9）进行实施。

$$h_r' = \delta\left(W_3 \hat{h}_r\right) \quad (9)$$

式中，h_r'表示第r个图节点的处理后特征；$\hat{h}_r$表示第r个图节点的近邻融合特征；W_3属于第m_2个图注意力层的层参数；$\delta(\ )$表示sigmoid函数。

基于上述步骤71至步骤74的相关内容可知，对于第m_2个图注意力层来说，该第m_2个图注意力层可以借助公式（7）-（9），对该第m_2个图注意力层的输入数据进行图注意力处理，得到该第m_2个图注意力层的输出数据。

实际上，因待使用图中各个候选语句节点对应于两种连接边（例如，近邻表征边以及相关表征边），故为了更好地进行图注意力处理，本申请实施例还提供了确定上述“第r个图节点的处理后特征”的另一种可能的实施方式，在该实施方式中，该“第r个图节点的处理后特征”的确定过程不仅包括上述步骤71-步骤73，可以还包括步骤75-步骤77：

步骤75：确定第r个图节点的融合信息差特征。

上述“第r个图节点的融合信息差特征”用于表示该第r个图节点针对待使用用户记录的语义补充信息。

另外，本申请实施例不限定步骤75，例如，具体可以包括步骤751-步骤755：

步骤751：判断第r个图节点是否属于候选语句节点，若是，则执行步骤752-步骤754；若否，则执行步骤755。

本申请实施例中，对于第m_2个图注意力层来说，在该第m_2个图注意力层获取到第r个图节点的处理前特征之后，可以判断该第r个图节点是否属于候选语句节点；若该第r个图节点属于候选语句节点，则可以确定不仅存在该第r个图节点的近邻表征边，还存在该第r个图节点的相关表征边，故可以综合这两种连接边，对该第r个图节点进行图注意力处理；若该第r个图节点不属于候选语句节点，则可以确定该第r个图节点属于记录语句节点，从而可以确定只存在该第r个图节点的近邻表征边，故可以只参考该第r个图节点的近邻表征边，对该第r个图节点进行图注意力处理即可。

步骤752：根据第r个图节点的处理前特征、以及第g个记录语句节点的处理前特征，确定该第r个图节点对应的第g个记录语句注意力权重特征。其中，g为正整数，g≤G，G为正整数。

上述“第r个图节点对应的第g个记录语句注意力权重特征”用于表示该第r个图节点与

该第g个记录语句节点之间的语义交互结果。

另外，本申请实施例不限定上述“第r个图节点对应的第g个记录语句注意力权重特征”的确定过程，例如，其可以利用公式（10）进行实施。

$$\beta_g^r = \frac{\exp(ReLu(\theta_2{}^T[W_4 h_r || W_4 h_g]))}{\sum_{z=1}^{G}\exp(ReLu(\theta_2{}^T[W_4 h_r || W_4 h_z]))} \quad (10)$$

式中，β_g^r表示第r个图节点对应的第g个记录语句注意力权重特征；h_r表示第r个图节点的处理前特征；h_g表示第g个记录语句节点的处理前特征；h_z表示第z个记录语句节点的处理前特征，且为z正整数，$z \leqslant G$；W_4、θ_2均属于第m_2个图注意力层的层参数，且这些层参数可以在上述“补充语句确定模型”的构建过程中确定；$ReLu(\,)$表示激活函数。

需要说明的是，本申请实施例不限定上述公式（10）的实现方式，例如，可以借助注意力层、激活函数层以及归一化处理层进行实现。

步骤753：根据G个记录语句节点的处理前特征、以及第r个图节点对应的G个记录语句注意力权重特征，确定该第r个图节点的记录融合特征。

需要说明的是，本申请实施例不限定步骤753的实施方式，例如，其具体可以采用公式（11）进行实施。

$$\hat{e}_r = \delta\left(\sum_{g=1}^{G}\beta_g^r W_5 h_g\right) \quad (11)$$

式中，$\hat{e}_r$表示第r个图节点的记录融合特征；β_g^r表示第r个图节点对应的第g个记录语句注意力权重特征；h_g表示第g个记录语句节点的处理前特征；W_5属于第m_2个图注意力层的层参数；$\delta(\,)$表示sigmoid函数。

步骤754：根据第r个图节点的近邻融合特征、以及该第r个图节点的记录融合特征，确定该第r个图节点的融合信息差特征。

需要说明的是，本申请实施例不限定步骤754的实施方式，例如，其具体可以采用公式（12）进行实施。

$$Differ_r = \hat{h}_r - \hat{e}_r \quad (12)$$

式中，$Differ_r$表示第r个图节点的融合信息差特征；$\hat{h}_r$表示第r个图节点的近邻融合特征；$\hat{e}_r$表示第r个图节点的记录融合特征。

步骤755：将预设零向量，确定为该第r个图节点的融合信息差特征。

上述“预设零向量”是一个全零向量；而且本申请实施例不限定该“预设零向量”的数据维度，例如，该“预设零向量”的数据维度与上文$(\hat{h}_r - \hat{e}_r)$的数据维度保持一致，以保证每个记录语句节点的特征表示与每个候选语句节点的特征表示始终保持相同数据维度。

基于上述步骤75的相关内容可知，对于第m_2个图注意力层来说，在该第m_2个图注意力层获取到第r个图节点的处理前特征之后，该第m_2个图注意力层可以根据该第r个图节点所具有的连接边类型，确定该第r个图节点的融合信息差特征。

步骤76：将第r个图节点的近邻融合特征与该第r个图节点的融合信息差特征进行拼接，

得到该第r个图节点的融合拼接特征。

步骤77：根据第r个图节点的融合拼接特征，确定该第r个图节点的处理后特征。

本申请实施例不限定步骤77的实施方式，例如，其可以采用公式（13）进行实施。

$$h'_r = \delta\left(W_6(\hat{h}_r||Differ_r)\right) \quad (13)$$

式中，h'_r表示第r个图节点的处理后特征；$\hat{h}_r$表示第r个图节点的近邻融合特征；$Differ_r$表示第r个图节点的融合信息差特征；$||$表示拼接处理；$(\hat{h}_r||Differ_r)$表示第r个图节点的融合拼接特征；W_6属于第m_2个图注意力层的层参数；$\delta(\ )$表示sigmoid函数。

基于上述步骤75至步骤77的相关内容可知，对于第m_2个图注意力层来说，该第m_2个图注意力层可以借助公式（7）-（8）、以及（10）-（13），对该第m_2个图注意力层的输入数据进行图注意力处理，得到该第m_2个图注意力层的输出数据。

基于上述“图注意力模块”的相关内容可知，该“图注意力模块”可以包括多个图注意力层，而且每个图注意力层都可以借助公式（7）-（8）、以及（10）-（13）实现图注意力处理过程。

基于上述步骤62的相关内容可知，对于补充语句确定模型来说，在获取到待使用图中各个图节点的待使用特征之后，可以由该补充语句确定模型中图注意力模块针对这些图节点的待使用特征进行图注意力处理，得到并输出各个图节点的图注意特征。

步骤63：将至少一个图节点的图注意特征输入分类预测层，得到该分类预测层输出的关键语句预测结果。

上述“分类预测层”用于针对该分类预测层的输入数据进行二分类预测处理（也就是，判断各个待使用语句是否属于关键语句）；而且本申请实施例不限定该“分类预测层”的实施方式，例如，可以采用现有的或者未来出现的任意一种分类网络进行实施。

基于上述步骤61至步骤63的相关内容可知，对于包括图注意力模块、以及分类预测层的补充语句确定模型来说，该补充语句确定模型可以按照图5所示的关键语句抽取过程，对上文“至少一个待使用句子”进行关键补充语句确定处理，得到并输出关键语句预测结果。

另外，上述“补充语句确定模型”可以预先根据重点语句标注语料进行构建；而且本申请实施例不限定上述“补充语句确定模型”的构建过程，例如，其可以采用现有的或者未来出现的任意一种模型构建方法进行实施。

上述“重点语句标注语料”包括第二会议文本句子集合、该第二会议文本句子集合对应的样本记录语句、该第二会议文本句子集合对应的重点补充语句标签。

上述“第二会议文本句子集合”是通过针对样本会议的会议转写文本进行句子切分得到的；而且本申请实施例不限定该“第二会议文本句子集合”与上文“第一会议文本句子集合”之间的关系，例如，这两者可以相同，也可以不同。

上述“第二会议文本句子集合对应的样本记录语句”用于表示该第二会议文本句子集合对应的用户记录携带语义信息；而且该“第二会议文本句子集合对应的样本记录语句”属于该第二会议文本句子集合。需要说明的是，本申请实施例不限定上述“第二会议文本句子集合对应的样本记录语句”的获取过程，例如，可以由人工标注。又如，可以借助上文“语义匹配模型”进行确定。

上述“第二会议文本句子集合对应的重点补充语句标签”用于表示该第二会议文本句子集合中实际的重点补充语句；而且本申请实施例不限定该“第二会议文本句子集合对应的重点补充语句标签”，例如，其可以表示为{（s_1，lab_1），（s_2，lab_2），……，（s_U，lab_U）}。其中，s_u表示第二会议文本句子集合中的第 u 个语句。lab_u用于表示第二会议文本句子集合中的第 u 个语句是否实际属于重点补充语句，而且 lab_u=1，则表示该第 u 个语句实际属于重点补充语句，若 lab_u=0，则表示该第 u 个语句实际不属于重点补充语句。

基于上述 S4212 的相关内容可知，在获取到待使用图之后，可以借助预先构建的补充语句确定模型，针对该待使用图进行关键语句确定处理，得到并输出该待使用图对应的关键语句预测结果，以使该关键语句预测结果能够表示出各个待使用句子属于关键补充语句的可能性（尤其是，各个非记录语句属于关键补充语句的可能性）。

S4213：根据关键语句预测结果，确定至少一个关键补充语句。

作为示例，当上述“关键语句预测结果”包括各个待使用句子的补充语句得分时，S4213 具体可以为：按照各个待使用句子的补充语句得分，从所有待使用句子中筛选出满足预设关键补充条件的各个关键补充语句。

上述“预设关键补充条件”可以预先设定，例如，其具体可以为：补充语句得分高于预先设定的第七得分阈值。又如，在将所有待使用句子按照补充语句得分从大到小进行排序之后，该“预设关键补充条件”具体可以为：排列序号低于预先设定的第五序号阈值。

基于上述 S4213 的相关内容可知，在获取到关键语句预测结果之后，可以参考该关键语句预测结果所记录的补充语句得分，确定出一些补充语句得分比较高的待使用句子，均确定为关键补充语句，以使该关键补充语句能够更好地表示出针对上文“待使用用户记录”的语义补充信息。

基于上述 S421 的相关内容可知，在获取到待使用用户记录对应的至少一个目标句子之后，可以由补充语句确定模型参考这些目标句子，从所有待使用句子中确定出一些关键补充语句，以使这些关键补充语句能够表示出针对上文“待使用用户记录”的语义补充信息，以便后续能够基于这些关键补充语句进行纪要内容确定处理。

S422：根据待使用用户记录、至少一个目标句子、以及至少一个关键补充语句，确定待使用纪要内容。

需要说明的是，本申请实施例不限定 S422 的实施方式，例如，如图 2 所示，S422 具体可以为：可以将这些关键补充语句、目标句子以及待使用用户记录均显示给用户，以使该用户能够基于这些显示内容编写待使用纪要内容，从而使得该用户能够从这些显示内容中获知其编写纪要所需的比较完成的会议记录内容，如此能够有效地避免因用户通过重听会议录制音频确定与该待使用用户记录相关的会议内容而导致的时长消耗，从而能够有效地提高纪要确定效率。

还需要说明的是，本申请实施例不限定上述“至少一个关键补充语句”的显示触发条件，例如，可以响应于用户触发的关键补充语句显示请求进行显示。其中，“关键补充语句显示请求”用于请求显示针对用户记录的重点补充语句；而且本申请实施例不限定该“关键补充语句显示请求”的触发方式，例如，用户可以通过点击图 2 所示的“查缺补漏”按钮进行触发该“关键补充语句显示请求”。

另外，为了进一步提高纪要确定效果，S422 可以采用下文**方法实施例六**所示的 S422 的任一实施方式进行实施。

基于上述 S421 至 S422 的相关内容可知，在获取到待使用用户记录对应的至少一个目标句子之后，可以先参考这些目标句子，从待使用句子中确定出一些关键补充语句，以使这些关键补充语句能够表示出针对上文“待使用用户记录”的语义补充信息；再参考这些关键补充语句、这些目标句子、以及待使用用户记录，确定待使用纪要内容。

方法实施例六

实际上，为了进一步提高纪要确定效果，可以从关键补充语句中提取要点提醒语句，使得该要点提醒语句能够以更简洁的内容描述出该关键补充语句携带的语义信息。基于此，本申请实施例还提供了 S422 的另一种可能的实施方式，其具体可以包括 S4221- S4222：

S4221：对各关键补充语句进行要点抽取处理，得到各关键补充语句的要点内容。

其中，“第 f 个关键补充语句的要点内容”是指该第 f 个关键补充语句的核心语义信息。其中，f 为正整数，f≤F，F 为正整数，F 表示关键补充语句的个数。

另外，本申请实施例不限定上述“第 f 个关键补充语句的要点内容”的确定过程，例如，可以采用现有的或者未来出现的任意一种语句核心抽取方法（例如，摘要抽取方法、句子简化方法等）进行实施。

又如，为了提高要点内容抽取效果，本申请实施例还提供了确定上述“第 f 个关键补充语句的要点内容”的另一种可能的实施方式，其具体可以包括：根据该第 f 个关键补充语句和预先构建的要点抽取模型，确定该第 f 个关键补充语句的要点内容。

上述“要点抽取模型”用于针对该要点抽取模型的输入数据进行要点内容抽取处理；而且该“要点抽取模型”可以是一种机器学习模型。

另外，本申请实施例不限定上述“要点抽取模型”，例如，其可以是基于编码-解码（Encoder–Decoder，ED）的文本生成模型。

此外，本申请实施例不限定上述“第 f 个关键补充语句的要点内容”的确定过程，例如，其具体可以包括：将第 f 个关键补充语句输入预先构建的要点抽取模型，得到该要点抽取模型输出的该第 f 个关键补充语句的要点内容。其中，要点抽取模型是根据大量第一样本语句、以及该第一样本语句对应的实际要点内容进行构建的；而且该“第一样本语句对应的实际要点内容”可以采用人工标注的方式进行确定。

实际上，上述“关键补充语句”可能存在信息缺失的现象（例如，指代省略等信息不完整现象），故为了提高提高要点内容抽取效果，可以参考该“关键补充语句”的上下文信息，对该“关键补充语句”进行要点内容抽取处理。基于此，本申请实施例还提供了确定上述“第 f 个关键补充语句的要点内容”的另一种可能的实施方式，其具体可以包括 S42211-S42212：

S42211：根据至少一个待使用句子，确定第 f 个关键补充语句的上下文信息。

上述“第 f 个关键补充语句的上下文信息”用于描述该第 f 个关键补充语句的上下文语义信息；而且本申请实施例不限定该“第 f 个关键补充语句的上下文信息”，例如，其可以包括一个或者多个待使用句子。

另外，本申请实施例不限定上述“第 f 个关键补充语句的上下文信息”的确定过程，例如，其具体可以为：将所有待使用句子，确定为第 f 个关键补充语句的上下文信息。

又如，上述“第 f 个关键补充语句的上下文信息”的确定过程，具体可以为：按照第 f 个关键补充语句的文本位置以及预设上下文窗口，从上述“至少一个待使用句子”中抽取该第 f 个关键补充语句的上下文信息，以使该“第 f 个关键补充语句的上下文信息”中所有语句均位于该预设上下文窗口内。其中，“预设上下文窗口”可以预先设定。

S42212：将第 f 个关键补充语句、以及该第 f 个关键补充语句的上下文信息输入要点抽取模型，得到该要点抽取模型输出的该第 f 个关键补充语句的要点内容。

上述“要点抽取模型”可以根据大量第二样本语句、该第二样本语句的上下文信息以及该第二样本语句对应的实际要点内容进行构建。其中，“第二样本语句对应的实际要点内容”可以采用人工标注的方式进行确定。

另外，本申请实施例不限定上述“要点抽取模型”的网络结构，例如，可以采用现有的或者未来出现的任意一种文本生成模型进行实施。

实际上，距离第 f 个关键补充语句越近的句子越有可能携带有该第 f 个关键补充语句所缺失的语义信息，故为了进一步提高要点内容提取效果，可以参考该第 f 个关键补充语句的上下文信息中各个语句的文本位置，对该第 f 个关键补充语句进行要点内容抽取处理。基于此，本申请实施例还提供了上述“要点抽取模型”的一种可能的实施方式，其具体可以包括：编码器、位置加权模块、注意力模块、融合模块以及解码器。

为了便于理解上述“要点抽取模型”的工作原理，下面以上述“第 f 个关键补充语句的上下文信息”的确定过程为例进行说明。

作为示例，当上述“第 f 个关键补充语句的上下文信息”包括 K 个参考语句时，该“第 f 个关键补充语句的上下文信息”的确定过程，具体可以包括步骤 81-步骤 85：

步骤 81：利用编码器，对第 f 个关键补充语句、以及各参考语句分别进行编码处理，得到该第 f 个关键补充语句的编码特征、以及各参考语句的编码特征。

上述“编码器”用于针对该编码器的输入数据进行编码处理；而且本申请实施例不限定该“编码器”，例如，其可以采用现有的或者未来出现的任意一种编码网络（例如，BERT 预训练模型等）进行实施。

上述“第 f 个关键补充语句的编码特征”用于表示该第 f 个关键补充语句携带的语义信息；而且该“第 f 个关键补充语句的编码特征”的确定过程具体可以为：将该第 f 个关键补充语句输入编码器，得到该编码器输出的该第 f 个关键补充语句的编码特征。

“第 k 个参考语句的编码特征”用于表示该第 k 个参考语句携带的语义信息；而且该“第 k 个参考语句的编码特征”的确定过程具体可以为：利用编码器对该第 k 个参考语句进行编码处理，得到该编码器输出的该第 k 个参考语句的编码特征。其中，k 为正整数，k ≤K，K 为正整数。

步骤 82：将 K 个参考语句的编码特征输入位置加权模块，得到该位置加权模块输出的该 K 个参考语句的位置加权特征。

其中，“第 k 个参考语句的位置加权特征”用于表示该第 k 个参考语句携带的语义信息以及文本位置影响信息。

上述“位置加权模块”用于针对该位置加权模块的输入数据进行位置加权处理；而且本申请实施例不限定该“位置加权模块”的工作原理，为了便于理解，下面以上述“第 k 个参考语句的位置加权特征”的确定过程为例进行说明。

作为示例，当该“位置加权模块”的模块参数包括 K 个位置加权向量时，上述“第 k 个参考语句的位置加权特征”的确定过程，具体可以包括步骤 821-步骤 822：

步骤 821：根据第 k 个位置加权向量与第 k 个参考语句的编码特征之间的点积，确定该第 k 个参考语句对应的位置激励权重。

上述“第 k 个位置加权向量”用于表示第 k 个参考语句的文本位置对上述“第 f 个关键补充语句”的影响程度；而且该“第 k 个位置加权向量”属于上述“位置加权模块”的模块参数，使得该“第 k 个位置加权向量”可以在上述“要点抽取模型”的构建过程中确定。

另外，本申请实施例不限定该“第 k 个位置加权向量”的数据维度，例如，该“第 k 个位置加权向量”的数据维度与上文“第 k 个参考语句的编码特征”的数据维度保持一致。

上述“第 k 个参考语句对应的位置激励权重”用于表示该第 k 个参考语句的携带的语义信息对上述“第 f 个关键补充语句”的信息补充程度；而且，本申请实施例不限定该“第 k 个参考语句对应的位置激励权重”的确定过程，例如，其可以采用公式（14）进行实施。

$$u_k = \delta\left(\mathrm{POSMatrix}_k{}^T h_{f,[CLS]}\right) \tag{14}$$

式中，u_k表示第 k 个参考语句对应的位置激励权重；$h_{f,[CLS]}$表示字符串[第 k 个参考语句，预设句子标识符]的编码特征；POSMatrix表示位置加权矩阵，该POSMatrix包括 K 个位置加权向量，且该POSMatrix属于上述“位置加权模块”的模块参数，使得该POSMatrix可以在上述“要点抽取模型”的构建过程中确定；$\mathrm{POSMatrix}_k$表示第 k 个位置加权向量；$\mathrm{POSMatrix}_k{}^T$表示第 k 个位置加权向量的转置；δ( )表示 sigmoid 函数。

步骤 822：将第 k 个参考语句对应的位置激励权重与该第 k 个参考语句的编码特征之间的乘积，确定为该第 k 个参考语句的位置加权特征（如公式（15）所示）。

$$\hat{h}_k = u_k * h_f = \delta\left(\mathrm{POSMatrix}_k{}^T h_{f,[CLS]}\right) * h_f \tag{15}$$

式中，$\hat{h}_k$表示第 k 个参考语句的位置加权特征；u_k表示第 k 个参考语句对应的位置激励权重；h_f表示第 k 个参考语句的编码特征；$h_{f,[CLS]}$表示字符串[第 k 个参考语句，预设句子标识符]的编码特征；POSMatrix表示位置加权矩阵，该POSMatrix包括 K 个位置加权向量，且该POSMatrix属于上述“位置加权模块”的模块参数，使得该POSMatrix可以在上述“要点抽取模型”的构建过程中确定；$\mathrm{POSMatrix}_k$表示第 k 个位置加权向量；$\mathrm{POSMatrix}_k{}^T$表示第 k 个位置加权向量的转置；δ( )表示 sigmoid 函数。

基于上述步骤 82 的相关内容可知，对于要点抽取模型来说，在获取到第 k 个参考语句的编码特征之后，可以借助该要点抽取模型中位置加权模块，对该第 k 个参考语句的编码特征以及该第 k 个参考语句对应的位置加权向量进行融合处理，得到该第 k 个参考语句的

位置加权特征，以使该“第 k 个参考语句的位置加权特征”能够更好地表示出该第 k 个参考语句携带的语义补充信息。

步骤 83：将第 k 个参考语句的位置加权特征、以及第 f 个关键补充语句的编码特征输入注意力模块，得到该注意力模块输出的第 f 个关键补充语句对应的第 k 个融合特征。其中，k 为正整数，k≤K，K 为正整数。

上述“注意力模块”用于针对该注意力模块的输入数据进行注意力处理；而且本申请实施例不限定该“注意力模块”的实施方式，例如，可以采用多头注意力网络进行实施。

上述“第 f 个关键补充语句对应的第 k 个融合特征”用于表示第 k 个参考语句与该第 f 个关键补充语句的融合语义信息。

另外，本申请实施例不限定上述“第 f 个关键补充语句对应的第 k 个融合特征”的获取过程，例如，当上述“注意力模块”采用多头注意力网络进行实施时，可以将第 f 个关键补充语句的编码特征，作为该多头注意力网络所需使用的询问参数（也就是，Q 参数），并且将第 k 个参考语句的位置加权特征，作为该多头注意力网络所需使用的其他两个参数（也就是，K 参数和 V 参数），以使该多头注意力网络能够通过多头注意力机制实现信息融合，得到该“第 f 个关键补充语句对应的第 k 个融合特征”。

步骤 84：将第 f 个关键补充语句对应的 K 个融合特征输入融合模块，得到该融合模块输出的该第 f 个关键补充语句的表征特征。

上述“融合模块”用于针对该融合模块的输入数据进行融合处理；而且本申请实施例不限定该“融合模块”的实施方式，例如，可以采用最大池化（max pooling）方法进行实施。

上述“第 f 个关键补充语句的表征特征”用于表示该第 f 个关键补充语句与 K 个参考语句的语义融合信息。

另外，本申请实施例不限定上述“第 f 个关键补充语句的表征特征”，例如，如图 6 所示，该“第 f 个关键补充语句的表征特征”具体可以为：将第 f 个关键补充语句对应的第 1 个融合特征、该第 f 个关键补充语句对应的第 2 个融合特征、……、以及该第 f 个关键补充语句对应的第 K 个融合特征进行最大池化处理，得到该“第 f 个关键补充语句的表征特征”，以使该“第 f 个关键补充语句的表征特征”能够更准确地表示出该第 f 个关键补充语句携带的语义信息。

步骤 85：将第 f 个关键补充语句的表征特征输入解码器，得到该解码器输出的该第 f 个关键补充语句的要点内容。

上述“解码器”用于针对该解码器的输入数据进行解码处理；而且本申请实施例不限定上述“解码器”的实施方式，例如，其可以采用现有的或者未来出现的任意一种解码网络（也就是，Decoder）进行实施。

基于上述步骤 81 至步骤 85 的相关内容可知，对于包括编码器、位置加权模块、注意力模块、融合模块以及解码器的要点抽取模型来说，其可以按照图 6 所示的要点内容生成过程进行实施，以使该要点抽取模型能够准确地确定出第 f 个关键补充语句的要点内容。

需要说明的是，对于图 6 来说，s_f表示第 f 个关键补充语句；$s_{k\in win,k\neq f}$表示第 k 个参考语句；$\hat{h}_k$表示第 k 个参考语句的位置加权特征，k∈{1，2，3，……，K}；$\mathrm{POSMatrix}_k$表

示第 k 个位置加权向量；$H_{f,k}$表示第 f 个关键补充语句对应的第 k 个融合特征，k∈{1，2，3，……，K}；h'_f表示第 f 个关键补充语句的表征特征；Y_f表示第 f 个关键补充语句的要点内容。

基于上述 S4221 的相关内容可知，在获取到第 f 个关键补充语句之后，可以先确定该第 f 个关键补充语句的上下文信息；再由预先构建的要点抽取模型，参考该第 f 个关键补充语句的上下文信息，对该第 f 个关键补充语句进行要点内容抽取处理，得到并输出该第 f 个关键补充语句的要点内容，以使该“第 f 个关键补充语句的要点内容”能够更准确地表示出该第 f 个关键补充语句的核心语义信息。其中，f 为正整数，f≤F，F 为正整数，F 表示关键补充语句的个数。

S4222：根据待使用用户记录、至少一个目标句子、至少一个关键补充语句、以及至少一个关键补充语句的要点内容，确定待使用纪要内容。

需要说明的是，本申请实施例不限定 S4222 的实施方式，例如，如图 2 所示，S4222 具体可以为：可以将这些关键补充语句及其要点内容、目标句子以及待使用用户记录均显示给用户，以使该用户能够基于这些显示内容编写待使用纪要内容，从而使得该用户能够从这些显示内容中获知其编写纪要所需的比较完成的会议记录内容，如此能够有效地避免因用户通过重听会议录制音频确定与该待使用用户记录相关的会议内容而导致的时长消耗，从而能够有效地提高纪要确定效率。

还需要说明的是，本申请实施例不限定上述“至少一个关键补充语句的要点内容”的显示触发条件，例如，可以响应于用户触发的要点补充显示请求进行显示。其中，“要点补充显示请求”用于请求显示针对用户记录的重点补充语句的要点内容；而且本申请实施例不限定该“要点补充显示请求”的触发方式，例如，用户可以通过点击图 2 所示的“查缺补漏”按钮进行触发该“要点补充显示请求”。

基于上述 S4221 至 S4222 的相关内容可知，在获取到针对待使用用户记录的至少一个关键补充语句之后，可以分别针对各个关键补充语句进行要点内容抽取处理，得到各个关键补充语句的要点内容，以使这些要点内容能够更简洁地表示出该待使用用户记录遗漏的重要信息，以便用户能够从这些要点内容中更快速地获知该待使用用户记录的要点补充信息，如此有利于提高纪要确定效率。

方法实施例七

实际上，为了提高纪要确定效果，可以借助人机交互过程实现纪要内容编写处理。基于此，本申请实施例提供了上文 S4 的另一种可能的实施方式，例如，其具体可以包括步骤 91-步骤 92：

步骤91：根据待使用用户记录和待使用语义匹配结果，确定提示信息。

上述“提示信息”是指在用户进行纪要编写过程中向该用户所提供的信息；而且本申请实施例不限定该“提示信息”，例如，其可以包括待使用用户记录、该待使用用户记录对应的至少一个目标句子、至少一个关键补充语句、以及该至少一个关键补充语句的要点内容中的至少一个。

基于上述步骤91的相关内容可知，在获取到待使用用户记录和待使用语义匹配结果之后，可以基于该待使用用户记录和待使用语义匹配结果，确定提示信息，以便后续用户能够基于该提示信息编写纪要内容。

步骤92：向用户显示提示信息，以使用户参考提示信息编写待使用纪要内容。

本申请实施例不限定上述“提示信息”的显示过程，例如，其具体可以包括（1）-（5）：

（1）在用户输入待使用用户记录（例如，图2所示的“价格合理，渠道操作性”或者“区分性定价，差异化”）之后，可以显示该待使用用户记录以及该待使用用户记录对应的定位触发按钮（例如，图2所示的望远镜）。

（2）在用户点击待使用用户记录对应的定位触发按钮之后，可以按照第一显示方式，将该待使用用户记录对应的至少一个目标句子在待使用记录文本中进行显示，以使该用户能够从该待使用记录文本中查看到按照第一显示方式的各个目标句子。其中，“第一显示方式”可以预先设定，例如，如图2所示，该“第一显示方式”可以包括：以第一颜色（例如，黄色）进行高亮显示，并显示这些目标句子与待使用用户记录之间的对应关系。

（3）在用户触发关键补充语句显示请求（例如，点击图2所示的“查缺补漏”按钮）之后，可以按照第二显示方式，将上述“至少一个关键补充语句” 在待使用记录文本中进行显示，以使该用户能够从该待使用记录文本中查看到按照第二显示方式的各个关键补充语句。其中，“第二显示方式”可以预先设定，例如，如图2所示，该“第二显示方式”可以包括：以第二颜色（例如，红色）进行高亮显示。

（4）在用户触发的要点补充显示请求（例如，点击图 2 所示的“查缺补漏”按钮）之后，可以按照第三显示方式，对各个关键补充语句的要点内容进行显示。其中，“第三显示方式”可以预先设定，例如，如图 2 所示，该“第三显示方式”可以包括：显示各个关键补充语句与各个关键补充语句的要点内容之间的对应关系。

基于上述步骤 91 至步骤 92 的相关内容可知，在获取到待使用用户记录和待使用语义匹配结果之后，可以先基于该待使用用户记录和该待使用语义匹配结果，确定提示信息（例如，图 2 所示的“待使用用户记录”、“目标句子”、“关键补充语句”、以及“关键补充语句的要点内容”等）；再将这些提示信息显示给用户，以使该用户能够参考这些提示信息编写待使用纪要内容，如此有利于提高纪要确定效果。

另外，为了进一步提高用户的纪要编写体验，可以根据用户实时输入的纪要内容，更新提示信息（例如，更新关键补充语句及其要点内容等）。基于此，本申请实施例还提供了纪要确定方法的另一种可能的实施方式，在该实施方式中，当上述“提示信息”包括至少一个关键补充语句时，该纪要确定方法除了包括上述全部或者部分步骤以外，在上述步骤 92 之后，该纪要确定方法可以还包括步骤 93-步骤 97：

步骤 93：获取用户输入的当前纪要编写语句。

上述“当前纪要编写语句”用于表示在用户编写纪要内容时实时输入的句子。

步骤 94：根据当前纪要编写语句与各关键补充语句之间的语义匹配表征数据，从至少一个关键补充语句中查找满足第五匹配条件的待删除补充语句，得到查找结果。

其中，“当前纪要编写语句与第 f 个关键补充语句之间的语义匹配表征数据”用于表示该当前纪要编写语句携带的语义信息与该第 f 个关键补充语句携带的语义信息之间的相似

程度。其中，f 为正整数，f≤F，F 为正整数，F 表示关键补充语句的个数。

另外，本申请实施例不限定上述“当前纪要编写语句与第 f 个关键补充语句之间的语义匹配表征数据”的确定过程，例如，可以借助现有的或者未来出现的任意一种语义匹配方法进行实施。又如，可以借助上文“语义匹配模型”进行实施。

上述“第五匹配条件”可以预先设定，例如，其具体可以为：当前纪要编写语句与待删除补充语句之间的语义匹配表征数据达到预先设定的预设语义匹配阈值。

上述“待删除补充语句”用于表示满足第五匹配条件的关键补充语句。

上述“查找结果”用于表示至少一个关键补充语句中是否存在待删除补充语句。

步骤 95：若查找结果表示至少一个关键补充语句中存在至少一个待删除补充语句，则从该至少一个关键补充语句中删除该至少一个待删除补充语句。

本申请实施例中，若查找结果表示至少一个关键补充语句中存在至少一个待删除补充语句，则可以确定当前纪要编写语句已经表达了该至少一个待删除补充语句所携带的语义信息，故为了避免这些待删除补充语句对后续纪要编写过程造成干扰，可以直接从该至少一个关键补充语句中删除该至少一个待删除补充语句，以使删除后的至少一个关键补充语句中不存在这些待删除补充语句，以便后续这些待删除补充语句不再作为提示信息显示给用户。

步骤 96：根据至少一个关键补充语句，更新提示信息。

需要说明的是，本申请实施例不限定步骤 96 的实施方式，为了便于理解，下面结合两种情况进行说明。

情况一，当上述“提示信息”包括至少一个关键补充语句，则步骤 96 具体可以为：在获取到删除后的至少一个关键补充语句之后，可以利用该删除后的至少一个关键补充语句，更新提示信息，以使更新后的提示信息不再包括上述“至少一个待删除补充语句”，从而使得这些待删除补充语句不再作为提示信息显示给用户，如此能够有效地避免这些待删除补充语句对后续纪要编写过程造成干扰。

情况二，当上述“提示信息”包括至少一个关键补充语句、以及该至少一个关键补充语句的要点内容，则步骤 96 具体可以为：根据至少一个关键补充语句以及该至少一个关键补充语句的要点内容，更新提示信息。

可见，在获取到删除后的至少一个关键补充语句之后，可以利用该删除后的至少一个关键补充语句、以及该删除后的至少一个关键补充语句的要点内容，更新提示信息，以使更新后的提示信息不再包括上述“至少一个待删除补充语句”、以及该“至少一个待删除补充语句” 的要点内容，从而使得这些待删除补充语句及其要点内容不再作为提示信息显示给用户，如此能够有效地避免这些待删除补充语句及其要点内容对后续纪要编写过程造成干扰。

步骤 97：判断是否达到预设停止条件，若是，则结束针对提示信息的更新过程；若否，则返回执行步骤 92 及其后续步骤。

上述“预设停止条件”可以预先设定，例如，其具体可以为：接收到用户触发的纪要编写完成请求，或者，所有的关键补充语句均已被删除。

基于上述步骤 91 至步骤 97 的相关内容可知，如图 7 所示，在用户不断编写纪要内容

的过程中，可以不断缩减显示给用户的提示信息（例如，不断缩减显示给用户的关键补充语句），以实时地向用户展示在其编写纪要内容时仍旧需要使用的信息，以便在该用户编写完成纪要内容（或者，所有关键补充语句均已被删除）时，停止针对提示信息的更新过程，如此有利于提高纪要编写效果。

基于上述方法实施例提供的纪要确定方法，本申请实施例还提供了一种纪要确定装置，下面结合附图进行解释和说明。

装置实施例

装置实施例对纪要确定装置进行介绍，相关内容请参见上述方法实施例。

参见图 8，该图为本申请实施例提供的一种纪要确定装置的结构示意图。

本申请实施例提供的纪要确定装置 800，包括：

信息获取单元 801，用于获取待使用用户记录和待使用记录文本；

文本划分单元 802，用于对所述待使用记录文本进行句子划分处理，得到至少一个待使用句子；

语义匹配单元 803，用于将所述待使用用户记录与所述至少一个待使用句子进行语义匹配处理，得到待使用语义匹配结果；

纪要确定单元 804，用于根据所述待使用用户记录和所述待使用语义匹配结果，确定待使用纪要内容。

在一种可能的实施方式中，所述语义匹配单元803，具体用于：根据所述待使用用户记录、所述至少一个待使用句子以及预先构建的语义匹配模型，确定所述待使用语义匹配结果；其中，所述语义匹配模型用于针对所述语义匹配模型的输入数据进行语义匹配处理。

在一种可能的实施方式中，所述语义匹配单元803，包括：

第一确定子单元，用于根据所述至少一个待使用句子以及预设句子标识符，确定待使用匹配文本；

匹配预测子单元，用于将所述待使用用户记录以及所述待使用匹配文本输入所述语义匹配模型，得到所述语义匹配模型输出的匹配预测结果；

第二确定子单元，用于根据所述匹配预测结果，确定所述待使用语义匹配结果。

在一种可能的实施方式中，所述语义匹配模型包括词编码模块、内容交互模块以及匹配预测模块；

所述匹配预测子单元，具体用于：根据所述待使用用户记录、所述待使用匹配文本以及词编码模块，确定用户记录编码特征和至少一个句子编码特征；根据所述用户记录编码特征、所述至少一个句子编码特征、以及所述内容交互模块，得到至少一个内容交互特征；将所述至少一个内容交互特征输入所述匹配预测模块，得到所述匹配预测模块输出的所述匹配预测结果。

在一种可能的实施方式中，所述语义匹配模型还包括句编码模块；

所述至少一个内容交互特征的确定过程，包括：将所述至少一个句子编码特征输入所述句编码模块，得到所述句编码模块输出的至少一个句子表征特征；根据所述用户记录编码特征、所述至少一个句子表征特征、以及所述内容交互模块，得到至少一个内容交互特

征。

在一种可能的实施方式中，所述至少一个内容交互特征的确定过程，包括：将所述用户记录编码特征以及所述至少一个句子表征特征输入所述内容交互模块，得到所述内容交互模块输出的所述至少一个内容交互特征。

在一种可能的实施方式中，所述至少一个内容交互特征的确定过程，包括：根据所述用户记录编码特征与各所述句子表征特征之间的相似度表征数据，从所述至少一个句子表征特征中筛选出满足第一相似条件的至少一个待参考句子特征；将所述用户记录编码特征以及所述至少一个待参考句子特征输入所述内容交互模块，得到所述内容交互模块输出的所述至少一个内容交互特征。

在一种可能的实施方式中，所述至少一个内容交互特征的确定过程，包括：根据所述用户记录编码特征与各所述句子编码特征之间的相似度表征数据，从所述至少一个句子编码特征中筛选出满足第二相似条件的至少一个待处理句子特征；将所述用户记录编码特征以及所述至少一个待处理句子特征输入所述内容交互模块，得到所述内容交互模块输出的所述至少一个内容交互特征。

在一种可能的实施方式中，所述内容交互模块包括注意力层和融合层；

当所述内容交互模块的输入数据包括所述用户记录编码特征以及至少一个待交互句子特征时，所述至少一个内容交互特征的确定过程，包括：将所述用户记录编码特征以及所述至少一个待交互句子特征输入所述注意力层，得到所述注意力层输出的用户记录表征特征；根据所述用户记录表征特征、所述至少一个待交互句子特征、以及所述融合层，确定所述至少一个内容交互特征。

在一种可能的实施方式中，所述待交互句子特征的个数为E；第e个内容交互特征的确定过程，包括：将所述用户记录表征特征与第e个待交互句子特征输入所述融合层，得到所述融合层输出的所述第e个内容交互特征；其中，e为正整数，e≤E，E为正整数。

在一种可能的实施方式中，所述第e个内容交互特征的确定过程，包括：根据所述用户记录表征特征和所述第e个待交互句子特征，确定第e个相似度表征特征、以及第e个信息差表征特征；其中，所述第e个相似度表征特征用于表示所述用户记录表征特征与所述第e个待交互句子特征之间的信息相似性；所述第e个信息差表征特征用于表示所述用户记录表征特征与所述第e个待交互句子特征之间的信息差异性；将所述用户记录表征特征、所述第e个待交互句子特征、所述第e个相似度表征特征、以及所述第e个信息差表征特征进行拼接处理，得到所述第e个内容交互特征。

在一种可能的实施方式中，所述第e个相似度表征特征是根据所述用户记录表征特征与所述第e个待交互句子特征之间的点积确定的。

在一种可能的实施方式中，所述第e个信息差表征特征是根据所述用户记录表征特征与所述第e个待交互句子特征之间的差值确定的。

在一种可能的实施方式中，所述第二确定子单元，具体用于：根据待使用文本检索结果、以及所述匹配预测结果，确定所述待使用语义匹配结果；其中，所述待使用文本检索结果是根据预设检索算法、所述待使用用户记录以及所述至少一个待使用句子确定的。

在一种可能的实施方式中，所述至少一个待使用句子包括至少一个语义匹配句子，且

所述待使用语义匹配结果包括所述待使用用户记录与各所述语义匹配句子之间的语义匹配得分；

所述纪要确定装置800，还包括：

匹配更新单元，用于根据所述待使用用户记录与各所述语义匹配句子之间的语义匹配得分，从所述至少一个语义匹配句子中筛选出满足预设基准条件的至少一个基准句子；根据所述至少一个基准句子的文本位置、以及各所述语义匹配句子的文本位置，确定各所述语义匹配句子的距离惩罚得分；利用各所述语义匹配句子的距离惩罚得分，更新所述待使用用户记录与各所述语义匹配句子之间的语义匹配得分。

在一种可能的实施方式中，所述语义匹配句子的个数为E；其中，E为正整数；

第e个语义匹配句子的距离惩罚得分的确定过程，包括：根据各个基准句子的文本位置与所述第e个语义匹配句子的文本位置之间的距离，确定所述第e个语义匹配句子对应的距离集；对所述第e个语义匹配句子对应的距离集进行预设统计分析处理，得到所述第e个语义匹配句子对应的待使用距离；根据所述第e个语义匹配句子对应的待使用距离，确定所述第e个语义匹配句子的距离惩罚得分。

在一种可能的实施方式中，所述至少一个待使用句子包括至少一个语义匹配句子，且所述待使用语义匹配结果包括所述待使用用户记录与各所述语义匹配句子之间的语义匹配得分；

所述纪要确定单元804，包括：

第一筛选子单元，用于根据所述待使用用户记录与各所述语义匹配句子之间的语义匹配得分，从所述至少一个语义匹配句子中筛选出满足第一匹配条件的至少一个目标句子；

第三确定子单元，用于根据所述待使用用户记录和所述至少一个目标句子，确定待使用纪要内容。

在一种可能的实施方式中，所述第一筛选子单元，包括：

第二筛选子单元，用于根据所述待使用用户记录与各所述语义匹配句子之间的语义匹配得分，从所述至少一个语义匹配句子中筛选出满足第二匹配条件的至少一个初选句子；

第三筛选子单元，用于根据所述待使用用户记录与各所述初选句子之间的语义匹配得分，从所述至少一个初选句子中筛选出满足第三匹配条件的至少一个目标句子。

在一种可能的实施方式中，所述第一筛选子单元，还包括：

第四筛选子单元，用于根据各所述初选句子的非停用词个数，从所述至少一个初选句子中筛选出满足第四匹配条件的至少一个再选句子；

所述第三筛选子单元，具体用于：根据所述待使用用户记录与各所述再选句子之间的语义匹配得分，从所述至少一个再选句子中筛选出满足所述第三匹配条件的至少一个目标句子。

在一种可能的实施方式中，所述至少一个目标句子的确定过程包括：根据所述待使用用户记录与各候选句子之间的语义匹配得分，从至少一个候选句子中筛选出满足第一丢弃条件的待删除句子，得到第一筛选结果；根据所述第一筛选结果和所述至少一个候选句子，确定所述至少一个目标句子。

在一种可能的实施方式中，所述至少一个目标句子的确定过程包括：若所述第一筛选

结果表示所述至少一个候选句子中存在至少一个待删除句子，则根据所述至少一个待删除句子和所述至少一个候选句子，确定至少一个待保留句子；根据所述至少一个待保留句子，确定所述至少一个目标句子。

在一种可能的实施方式中，所述至少一个目标句子的确定过程包括：若所述至少一个候选句子满足预设校正条件，则从所述至少一个待删除句子中筛选出满足第一保留条件的至少一个待评估句子；对所述待使用用户记录与各所述待评估句子之间的语义匹配得分进行归一化处理，得到各所述待评估句子的归一化得分；根据各所述待评估句子的归一化得分，从所述至少一个待评估句子中筛选出满足第二保留条件的至少一个第一保留句子；从所述至少一个候选句子中删除所述至少一个待删除句子，得到至少一个第二保留句子；根据所述至少一个第一保留句子、以及所述至少一个第二保留句子，确定所述至少一个待保留句子。

在一种可能的实施方式中，所述第三确定子单元，包括：

第四确定子单元，用于根据所述至少一个目标句子、所述至少一个待使用句子、以及预先构建的补充语句确定模型，确定至少一个关键补充语句；其中，所述补充语句确定模型用于针对所述补充语句确定模型的输入数据进行关键补充语句确定处理；

第五确定子单元，用于根据所述待使用用户记录、所述至少一个目标句子、以及至少一个关键补充语句，确定待使用纪要内容。

在一种可能的实施方式中，所述第四确定子单元，包括：

图构建子单元，用于根据所述至少一个目标句子和所述至少一个待使用句子，构建待使用图；

补充预测子单元，用于将所述待使用图输入所述补充语句确定模型，得到所述补充语句确定模型输出的关键语句预测结果；

语句补充子单元，用于根据所述关键语句预测结果，确定所述至少一个关键补充语句。

在一种可能的实施方式中，所述图构建子单元，具体用于：根据所述至少一个目标句子和所述至少一个待使用句子，确定至少一个图节点；其中，所述至少一个图节点包括至少一个记录语句节点和至少一个候选语句节点；按照预设近邻描述参数，构建所述至少一个图节点对应的近邻表征边；将各所述候选语句节点分别与各所述记录语句节点进行相连，得到各所述候选语句节点对应的相关表征边；根据所述至少一个图节点、所述至少一个图节点对应的近邻表征边、以及所述至少一个候选语句节点对应的相关表征边，确定所述待使用图。

在一种可能的实施方式中，所述补充语句确定模型包括图注意力模块和分类预测层；

所述补充预测子单元，具体用于：确定所述待使用图中各个图节点的待使用特征；将所述至少一个图节点的待使用特征以及所述待使用图输入所述图注意力模块，得到所述图注意力模块输出的至少一个图节点的图注意特征；将所述至少一个图节点的图注意特征输入所述分类预测层，得到所述分类预测层输出的所述关键语句预测结果。

在一种可能的实施方式中，所述图注意力模块包括M个图注意力层；其中，第1个图注意力层的输入数据包括所述至少一个图节点的待使用特征以及所述待使用图；所述第m_1个图注意力层的输入数据包括第m_1-1个图注意力层的输出数据以及所述待使用图，m_1为正整

数，$2 \leqslant m_1 \leqslant M$，M为正整数。

在一种可能的实施方式中，所述待使用图包括所述至少一个图节点对应的近邻表征边；所述至少一个图节点包括G个记录语句节点，且所述图节点的个数为R；其中，R为正整数，G为正整数；

当第m_2个图注意力层的输入数据包括R个图节点的处理前特征以及所述待使用图，所述第m_2个图注意力层的输出数据包括R个图节点的处理后特征，且m_2为正整数，$1 \leqslant m_2 \leqslant M$时，第r个图节点的处理后特征确定过程包括：根据所述第r个图节点对应的近邻表征边，确定所述第r个图节点对应的至少一个近邻节点；其中，r为正整数，$r \leqslant R$；根据所述第r个图节点的处理前特征、以及所述第r个图节点对应的第d_r个近邻节点的处理前特征，确定所述第d_r个近邻节点的近邻注意力权重特征；其中，d_r为正整数，$d_r \leqslant D_r$，D_r为正整数，D_r表示所述第r个图节点对应的近邻节点个数；根据所述第r个图节点对应的D_r个近邻节点的处理前特征、以及所述D_r个近邻节点的近邻注意力权重特征，确定所述第r个图节点的近邻融合特征；根据所述第r个图节点的近邻融合特征，确定所述第r个图节点的处理后特征。

在一种可能的实施方式中，所述至少一个图节点还包括至少一个候选语句节点；所述第r个图节点的处理后特征的确定过程，包括：若所述第r个图节点属于候选语句节点，则根据所述第r个图节点的处理前特征、以及第g个记录语句节点的处理前特征，确定所述第r个图节点对应的第g个记录语句注意力权重特征；其中，g为正整数，$g \leqslant G$，G为正整数；根据G个记录语句节点的处理前特征、以及所述第r个图节点对应的G个记录语句注意力权重特征，确定所述第r个图节点的记录融合特征；根据所述第r个图节点的近邻融合特征、以及所述第r个图节点的记录融合特征，确定所述第r个图节点的融合信息差特征；将所述第r个图节点的近邻融合特征与所述第r个图节点的融合信息差特征进行拼接，得到所述第r个图节点的融合拼接特征；根据所述第r个图节点的融合拼接特征，确定所述第r个图节点的处理后特征。

在一种可能的实施方式中，所述第五确定子单元，包括：

要点抽取子单元，用于对各所述关键补充语句进行要点抽取处理，得到各所述关键补充语句的要点内容；

第六确定子单元，用于根据所述待使用用户记录、所述至少一个目标句子、所述至少一个关键补充语句、以及所述至少一个关键补充语句的要点内容，确定待使用纪要内容。

在一种可能的实施方式中，所述关键补充语句的个数为F，且第f个关键补充语句的要点内容是根据所述第f个关键补充语句和预先构建的要点抽取模型进行确定的；其中，所述要点抽取模型用于针对所述要点抽取模型的输入数据进行要点内容抽取处理；f为正整数，$f \leqslant F$，F为正整数。

在一种可能的实施方式中，所述要点抽取子单元，包括：

第七确定子单元，用于根据所述至少一个待使用句子，确定所述第f个关键补充语句的上下文信息；

要点生成子单元，用于将所述第f个关键补充语句、以及所述第f个关键补充语句的上下文信息输入所述要点抽取模型，得到所述要点抽取模型输出的所述第f个关键补充语句的要点内容。

在一种可能的实施方式中，所述上下文信息包括K个参考语句；所述要点抽取模型包括编码器、位置加权模块、注意力模块、融合模块以及解码器；

所述第f个关键补充语句的要点内容的确定过程，包括：利用所述编码器，对所述第f个关键补充语句、以及各所述参考语句分别进行编码处理，得到所述第f个关键补充语句的编码特征、以及各所述参考语句的编码特征；将所述K个参考语句的编码特征输入所述位置加权模块，得到所述位置加权模块输出的所述K个参考语句的位置加权特征；将所述第k个参考语句的位置加权特征、以及所述第f个关键补充语句的编码特征输入所述注意力模块，得到所述注意力模块输出的所述第f个关键补充语句对应的第k个融合特征；其中，k为正整数，k≤K，K为正整数；将所述第f个关键补充语句对应的K个融合特征输入所述融合模块，得到所述融合模块输出的所述第f个关键补充语句的表征特征；将所述第f个关键补充语句的表征特征输入所述解码器，得到所述解码器输出的所述第f个关键补充语句的要点内容。

在一种可能的实施方式中，所述位置加权模块的模块参数包括K个位置加权向量；所述第k个参考语句的位置加权特征的获取过程，包括：根据第k个位置加权向量与所述第k个参考语句的编码特征之间的点积，确定所述第k个参考语句对应的位置激励权重；将所述第k个参考语句对应的位置激励权重与所述第k个参考语句的编码特征之间的乘积，确定为所述第k个参考语句的位置加权特征。

在一种可能的实施方式中，所述纪要确定单元804，包括：

提示确定子单元，用于根据所述待使用用户记录和所述待使用语义匹配结果，确定提示信息；

提示显示子单元，用于向用户显示所述提示信息，以使所述用户参考所述提示信息编写所述待使用纪要内容。

在一种可能的实施方式中，所述纪要确定单元804还包括：

提示更新子单元，用于若所述提示信息包括所述至少一个关键补充语句，则获取所述用户输入的当前纪要编写语句；根据所述当前纪要编写语句与各所述关键补充语句之间的语义匹配表征数据，从所述至少一个关键补充语句中查找满足第五匹配条件的待删除补充语句，得到查找结果；若所述查找结果表示所述至少一个关键补充语句中存在至少一个待删除补充语句，则从所述至少一个关键补充语句中删除所述至少一个待删除补充语句；根据所述至少一个关键补充语句，更新所述提示信息，并返回所述提示显示子单元继续执行所述向用户显示所述提示信息的步骤，直至达到预设停止条件。

在一种可能的实施方式中，所述提示信息还包括所述至少一个关键补充语句的要点内容；所述提示信息的更新过程，包括：根据所述至少一个关键补充语句以及所述至少一个关键补充语句的要点内容，更新所述提示信息。

进一步地，本申请实施例还提供了一种设备，包括：处理器、存储器、系统总线；

所述处理器以及所述存储器通过所述系统总线相连；

所述存储器用于存储一个或多个程序，所述一个或多个程序包括指令，所述指令当被所述处理器执行时使所述处理器执行上述纪要确定方法的任一种实现方法。

进一步地，本申请实施例还提供了一种计算机可读存储介质，所述计算机可读存储介

质中存储有指令，当所述指令在终端设备上运行时，使得所述终端设备执行上述纪要确定方法的任一种实现方法。

进一步地，本申请实施例还提供了一种计算机程序产品，所述计算机程序产品在终端设备上运行时，使得所述终端设备执行上述纪要确定方法的任一种实现方法。

通过以上的实施方式的描述可知，本领域的技术人员可以清楚地了解到上述实施例方法中的全部或部分步骤可借助软件加必需的通用硬件平台的方式来实现。基于这样的理解，本申请的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来，该计算机软件产品可以存储在存储介质中，如 ROM/RAM、磁碟、光盘等，包括若干指令用以使得一台计算机设备（可以是个人计算机，服务器，或者诸如媒体网关等网络通信设备，等等）执行本申请各个实施例或者实施例的某些部分所述的方法。

需要说明的是，本说明书中各个实施例采用递进的方式描述，每个实施例重点说明的都是与其他实施例的不同之处，各个实施例之间相同相似部分互相参见即可。对于实施例公开的装置而言，由于其与实施例公开的方法相对应，所以描述的比较简单，相关之处参见方法部分说明即可。

还需要说明的是，在本文中，诸如第一和第二等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来，而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且，术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含，从而使得包括一系列要素的过程、方法、物品或者设备不仅包括那些要素，而且还包括没有明确列出的其他要素，或者是还包括为这种过程、方法、物品或者设备所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括所述要素的过程、方法、物品或者设备中还存在另外的相同要素。

对所公开的实施例的上述说明，使本领域专业技术人员能够实现或使用本申请。对这些实施例的多种修改对本领域的专业技术人员来说将是显而易见的，本文中所定义的一般原理可以在不脱离本申请的精神或范围的情况下，在其它实施例中实现。因此，本申请将不会被限制于本文所示的这些实施例，而是要符合与本文所公开的原理和新颖特点相一致的最宽的范围。

权 利 要 求

1、一种纪要确定方法，其特征在于，所述方法包括：

获取待使用用户记录和待使用记录文本；

对所述待使用记录文本进行句子划分处理，得到至少一个待使用句子；

将所述待使用用户记录与所述至少一个待使用句子进行语义匹配处理，得到待使用语义匹配结果；

根据所述待使用用户记录和所述待使用语义匹配结果，确定待使用纪要内容。

2、根据权利要求1所述的方法，其特征在于，所述将所述待使用用户记录与所述至少一个待使用句子进行语义匹配处理，得到待使用语义匹配结果，包括：

根据所述至少一个待使用句子以及预设句子标识符，确定待使用匹配文本；

将所述待使用用户记录以及所述待使用匹配文本输入所述语义匹配模型，得到所述语义匹配模型输出的匹配预测结果；

根据所述匹配预测结果，确定所述待使用语义匹配结果。

3、根据权利要求2所述的方法，其特征在于，所述语义匹配模型包括词编码模块、内容交互模块以及匹配预测模块；

所述匹配预测结果的确定过程，包括：

根据所述待使用用户记录、所述待使用匹配文本以及词编码模块，确定用户记录编码特征和至少一个句子编码特征；

根据所述用户记录编码特征、所述至少一个句子编码特征、以及所述内容交互模块，得到至少一个内容交互特征；

将所述至少一个内容交互特征输入所述匹配预测模块，得到所述匹配预测模块输出的所述匹配预测结果。

4、根据权利要求3所述的方法，其特征在于，所述语义匹配模型还包括句编码模块；

所述根据所述用户记录编码特征、所述至少一个句子编码特征、以及所述内容交互模块，得到至少一个内容交互特征，包括：

将所述至少一个句子编码特征输入所述句编码模块，得到所述句编码模块输出的至少一个句子表征特征；

根据所述用户记录编码特征、所述至少一个句子表征特征、以及所述内容交互模块，得到至少一个内容交互特征。

5、根据权利要求4所述的方法，其特征在于，所述根据所述用户记录编码特征、所述至少一个句子表征特征、以及所述内容交互模块，得到至少一个内容交互特征，包括：

将所述用户记录编码特征以及所述至少一个句子表征特征输入所述内容交互模块，得到所述内容交互模块输出的所述至少一个内容交互特征。

6、根据权利要求4所述的方法，其特征在于，所述根据所述用户记录编码特征、所述至少一个句子表征特征、以及所述内容交互模块，得到至少一个内容交互特征，包括：

根据所述用户记录编码特征与各所述句子表征特征之间的相似度表征数据，从所述至

少一个句子表征特征中筛选出满足第一相似条件的至少一个待参考句子特征；

将所述用户记录编码特征以及所述至少一个待参考句子特征输入所述内容交互模块，得到所述内容交互模块输出的所述至少一个内容交互特征。

7、根据权利要求3所述的方法，其特征在于，所述根据所述用户记录编码特征、所述至少一个句子编码特征、以及所述内容交互模块，得到至少一个内容交互特征，包括：

根据所述用户记录编码特征与各所述句子编码特征之间的相似度表征数据，从所述至少一个句子编码特征中筛选出满足第二相似条件的至少一个待处理句子特征；

将所述用户记录编码特征以及所述至少一个待处理句子特征输入所述内容交互模块，得到所述内容交互模块输出的所述至少一个内容交互特征。

8、根据权利要求5-7中任一项所述的方法，其特征在于，所述内容交互模块包括注意力层和融合层；

当所述内容交互模块的输入数据包括所述用户记录编码特征以及至少一个待交互句子特征时，所述至少一个内容交互特征的确定过程，包括：

将所述用户记录编码特征以及所述至少一个待交互句子特征输入所述注意力层，得到所述注意力层输出的用户记录表征特征；

根据所述用户记录表征特征、所述至少一个待交互句子特征、以及所述融合层，确定所述至少一个内容交互特征。

9、根据权利要求8所述的方法，其特征在于，所述待交互句子特征的个数为E；

第e个内容交互特征的确定过程，包括：

将所述用户记录表征特征与第e个待交互句子特征输入所述融合层，得到所述融合层输出的所述第e个内容交互特征；其中，e为正整数，e≤E，E为正整数。

10、根据权利要求2所述的方法，其特征在于，所述根据所述匹配预测结果，确定所述待使用语义匹配结果，包括：

根据待使用文本检索结果、以及所述匹配预测结果，确定所述待使用语义匹配结果；其中，所述待使用文本检索结果是根据预设检索算法、所述待使用用户记录以及所述至少一个待使用句子确定的。

11、根据权利要求1所述的方法，其特征在于，所述至少一个待使用句子包括至少一个语义匹配句子，且所述待使用语义匹配结果包括所述待使用用户记录与各所述语义匹配句子之间的语义匹配得分；

所述根据所述待使用用户记录和所述待使用语义匹配结果，确定待使用纪要内容之前，所述方法还包括：

根据所述待使用用户记录与各所述语义匹配句子之间的语义匹配得分，从所述至少一个语义匹配句子中筛选出满足预设基准条件的至少一个基准句子；

根据所述至少一个基准句子的文本位置、以及各所述语义匹配句子的文本位置，确定各所述语义匹配句子的距离惩罚得分；

利用各所述语义匹配句子的距离惩罚得分，更新所述待使用用户记录与各所述语义匹配句子之间的语义匹配得分。

12、根据权利要求11所述的方法，其特征在于，所述语义匹配句子的个数为E；其中，

E为正整数；

第e个语义匹配句子的距离惩罚得分的确定过程，包括：

根据各个基准句子的文本位置与所述第e个语义匹配句子的文本位置之间的距离，确定所述第e个语义匹配句子对应的距离集；

对所述第e个语义匹配句子对应的距离集进行预设统计分析处理，得到所述第e个语义匹配句子对应的待使用距离；

根据所述第e个语义匹配句子对应的待使用距离，确定所述第e个语义匹配句子的距离惩罚得分。

13、根据权利要求1所述的方法，其特征在于，所述至少一个待使用句子包括至少一个语义匹配句子，且所述待使用语义匹配结果包括所述待使用用户记录与各所述语义匹配句子之间的语义匹配得分；

所述根据所述待使用用户记录和所述待使用语义匹配结果，确定待使用纪要内容，包括：

根据所述待使用用户记录与各所述语义匹配句子之间的语义匹配得分，从所述至少一个语义匹配句子中筛选出满足第一匹配条件的至少一个目标句子；

根据所述待使用用户记录和所述至少一个目标句子，确定待使用纪要内容。

14、根据权利要求13所述的方法，其特征在于，所述根据所述待使用用户记录与各所述语义匹配句子之间的语义匹配得分，从所述至少一个语义匹配句子中筛选出满足第一匹配条件的至少一个目标句子，包括：

根据所述待使用用户记录与各所述语义匹配句子之间的语义匹配得分，从所述至少一个语义匹配句子中筛选出满足第二匹配条件的至少一个初选句子；

根据所述待使用用户记录与各所述初选句子之间的语义匹配得分，从所述至少一个初选句子中筛选出满足第三匹配条件的至少一个目标句子。

15、根据权利要求14所述的方法，其特征在于，所述方法还包括：

根据各所述初选句子的非停用词个数，从所述至少一个初选句子中筛选出满足第四匹配条件的至少一个再选句子；

所述根据所述待使用用户记录与各所述初选句子之间的语义匹配得分，从所述至少一个初选句子中筛选出满足第三匹配条件的至少一个目标句子，包括：

根据所述待使用用户记录与各所述再选句子之间的语义匹配得分，从所述至少一个再选句子中筛选出满足所述第三匹配条件的至少一个目标句子。

16、根据权利要求13-15中任一项所述的方法，其特征在于，所述至少一个目标句子的确定过程包括：

根据所述待使用用户记录与各候选句子之间的语义匹配得分，从至少一个候选句子中筛选出满足第一丢弃条件的待删除句子，得到第一筛选结果；

根据所述第一筛选结果和所述至少一个候选句子，确定所述至少一个目标句子。

17、根据权利要求16所述的方法，其特征在于，所述根据所述第一筛选结果和所述至少一个候选句子，确定所述至少一个目标句子，包括：

若所述第一筛选结果表示所述至少一个候选句子中存在至少一个待删除句子，则根据

所述至少一个待删除句子和所述至少一个候选句子，确定至少一个待保留句子；

根据所述至少一个待保留句子，确定所述至少一个目标句子。

18、根据权利要求17所述的方法，其特征在于，所述根据所述至少一个待删除句子和所述至少一个候选句子，确定至少一个待保留句子，包括：

若所述至少一个候选句子满足预设校正条件，则从所述至少一个待删除句子中筛选出满足第一保留条件的至少一个待评估句子；

对所述待使用用户记录与各所述待评估句子之间的语义匹配得分进行归一化处理，得到各所述待评估句子的归一化得分；

根据各所述待评估句子的归一化得分，从所述至少一个待评估句子中筛选出满足第二保留条件的至少一个第一保留句子；

从所述至少一个候选句子中删除所述至少一个待删除句子，得到至少一个第二保留句子；

根据所述至少一个第一保留句子、以及所述至少一个第二保留句子，确定所述至少一个待保留句子。

19、根据权利要求13所述的方法，其特征在于，所述根据所述待使用用户记录和所述至少一个目标句子，确定待使用纪要内容，包括：

根据所述至少一个目标句子、所述至少一个待使用句子、以及预先构建的补充语句确定模型，确定至少一个关键补充语句；其中，所述补充语句确定模型用于针对所述补充语句确定模型的输入数据进行关键补充语句确定处理；

根据所述待使用用户记录、所述至少一个目标句子、以及至少一个关键补充语句，确定待使用纪要内容。

20、根据权利要求19所述的方法，其特征在于，所述根据所述至少一个目标句子、所述至少一个待使用句子、以及预先构建的补充语句确定模型，确定至少一个关键补充语句，包括：

根据所述至少一个目标句子和所述至少一个待使用句子，构建待使用图；

将所述待使用图输入所述补充语句确定模型，得到所述补充语句确定模型输出的关键语句预测结果；

根据所述关键语句预测结果，确定所述至少一个关键补充语句。

21、根据权利要求20所述的方法，其特征在于，所述根据所述至少一个目标句子和所述至少一个待使用句子，构建待使用图，包括：

根据所述至少一个目标句子和所述至少一个待使用句子，确定至少一个图节点；其中，所述至少一个图节点包括至少一个记录语句节点和至少一个候选语句节点；

按照预设近邻描述参数，构建所述至少一个图节点对应的近邻表征边；

将各所述候选语句节点分别与各所述记录语句节点进行相连，得到各所述候选语句节点对应的相关表征边；

根据所述至少一个图节点、所述至少一个图节点对应的近邻表征边、以及所述至少一个候选语句节点对应的相关表征边，确定所述待使用图。

22、根据权利要求20所述的方法，其特征在于，所述补充语句确定模型包括图注意力

模块和分类预测层；

所述关键语句预测结果的确定过程，包括：

确定所述待使用图中各个图节点的待使用特征；

将所述至少一个图节点的待使用特征以及所述待使用图输入所述图注意力模块，得到所述图注意力模块输出的至少一个图节点的图注意特征；

将所述至少一个图节点的图注意特征输入所述分类预测层，得到所述分类预测层输出的所述关键语句预测结果。

23、根据权利要求19所述的方法，其特征在于，所述根据所述待使用用户记录、所述至少一个目标句子、以及至少一个关键补充语句，确定待使用纪要内容，包括：

对各所述关键补充语句进行要点抽取处理，得到各所述关键补充语句的要点内容；

根据所述待使用用户记录、所述至少一个目标句子、所述至少一个关键补充语句、以及所述至少一个关键补充语句的要点内容，确定待使用纪要内容。

24、根据权利要求23所述的方法，其特征在于，所述关键补充语句的个数为F；

所述第f个关键补充语句的要点内容的确定过程，包括：

根据所述至少一个待使用句子，确定所述第f个关键补充语句的上下文信息；其中，f为正整数，f≤F，F为正整数；

将所述第f个关键补充语句、以及所述第f个关键补充语句的上下文信息输入所述要点抽取模型，得到所述要点抽取模型输出的所述第f个关键补充语句的要点内容。

25、根据权利要求24所述的方法，其特征在于，所述上下文信息包括K个参考语句；所述要点抽取模型包括编码器、位置加权模块、注意力模块、融合模块以及解码器；

所述第f个关键补充语句的要点内容的确定过程，包括：

利用所述编码器，对所述第f个关键补充语句、以及各所述参考语句分别进行编码处理，得到所述第f个关键补充语句的编码特征、以及各所述参考语句的编码特征；

将所述K个参考语句的编码特征输入所述位置加权模块，得到所述位置加权模块输出的所述K个参考语句的位置加权特征；

将所述第k个参考语句的位置加权特征、以及所述第f个关键补充语句的编码特征输入所述注意力模块，得到所述注意力模块输出的所述第f个关键补充语句对应的第k个融合特征；其中，k为正整数，k≤K，K为正整数；

将所述第f个关键补充语句对应的K个融合特征输入所述融合模块，得到所述融合模块输出的所述第f个关键补充语句的表征特征；

将所述第f个关键补充语句的表征特征输入所述解码器，得到所述解码器输出的所述第f个关键补充语句的要点内容。

26、根据权利要求1所述的方法，其特征在于，所述根据所述待使用用户记录和所述待使用语义匹配结果，确定待使用纪要内容，包括：

根据所述待使用用户记录和所述待使用语义匹配结果，确定提示信息；

向用户显示所述提示信息，以使所述用户参考所述提示信息编写所述待使用纪要内容。

27、根据权利要求26所述的方法，其特征在于，所述向用户显示所述提示信息之后，所述方法还包括：

若所述提示信息包括所述至少一个关键补充语句，则获取所述用户输入的当前纪要编写语句；

根据所述当前纪要编写语句与各所述关键补充语句之间的语义匹配表征数据，从所述至少一个关键补充语句中查找满足第五匹配条件的待删除补充语句，得到查找结果；

若所述查找结果表示所述至少一个关键补充语句中存在至少一个待删除补充语句，则从所述至少一个关键补充语句中删除所述至少一个待删除补充语句；

根据所述至少一个关键补充语句，更新所述提示信息，并继续执行所述向用户显示所述提示信息的步骤，直至达到预设停止条件。

28、一种纪要确定装置，其特征在于，包括：

信息获取单元，用于获取待使用用户记录和待使用记录文本；

文本划分单元，用于对所述待使用记录文本进行句子划分处理，得到至少一个待使用句子；

语义匹配单元，用于将所述待使用用户记录与所述至少一个待使用句子进行语义匹配处理，得到待使用语义匹配结果；

纪要确定单元，用于根据所述待使用用户记录和所述待使用语义匹配结果，确定待使用纪要内容。

29、一种设备，其特征在于，所述设备包括：处理器、存储器、系统总线；

所述处理器以及所述存储器通过所述系统总线相连；

所述存储器用于存储一个或多个程序，所述一个或多个程序包括指令，所述指令当被所述处理器执行时使所述处理器执行权利要求1至27任一项所述的方法。

30、一种计算机可读存储介质，其特征在于，所述计算机可读存储介质中存储有指令，当所述指令在终端设备上运行时，使得所述终端设备执行权利要求1至27任一项所述的方法。

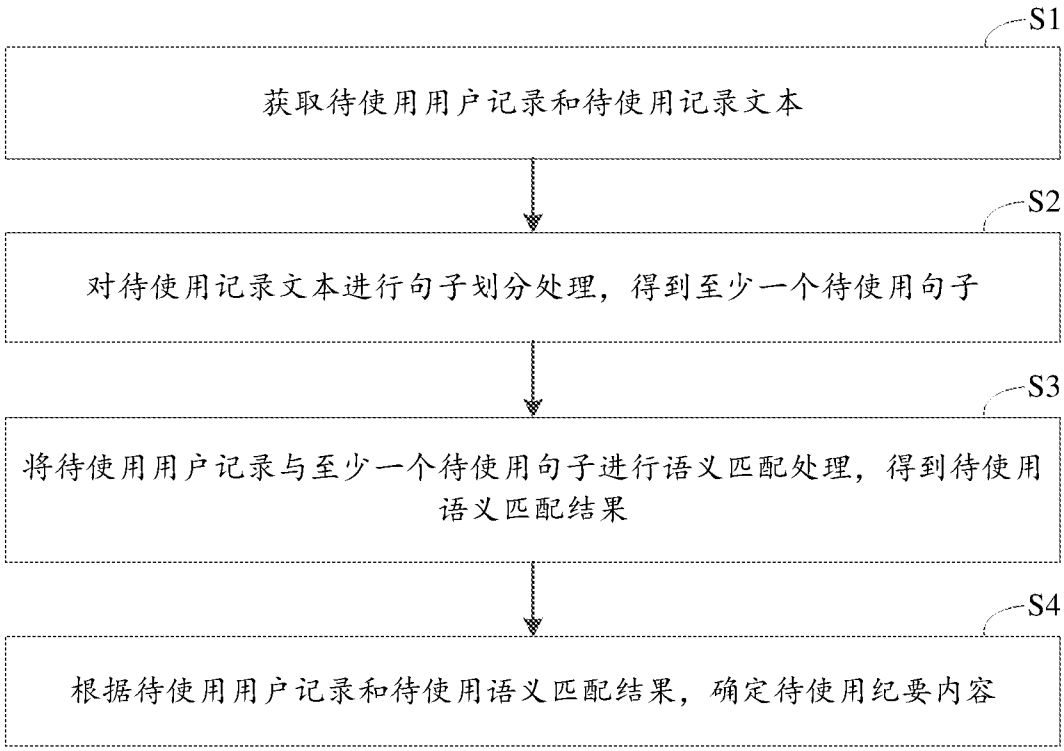

S1
获取待使用用户记录和待使用记录文本
S2
对待使用记录文本进行句子划分处理，得到至少一个待使用句子
S3
将待使用用户记录与至少一个待使用句子进行语义匹配处理，得到待使用语义匹配结果
S4
根据待使用用户记录和待使用语义匹配结果，确定待使用纪要内容

图 1

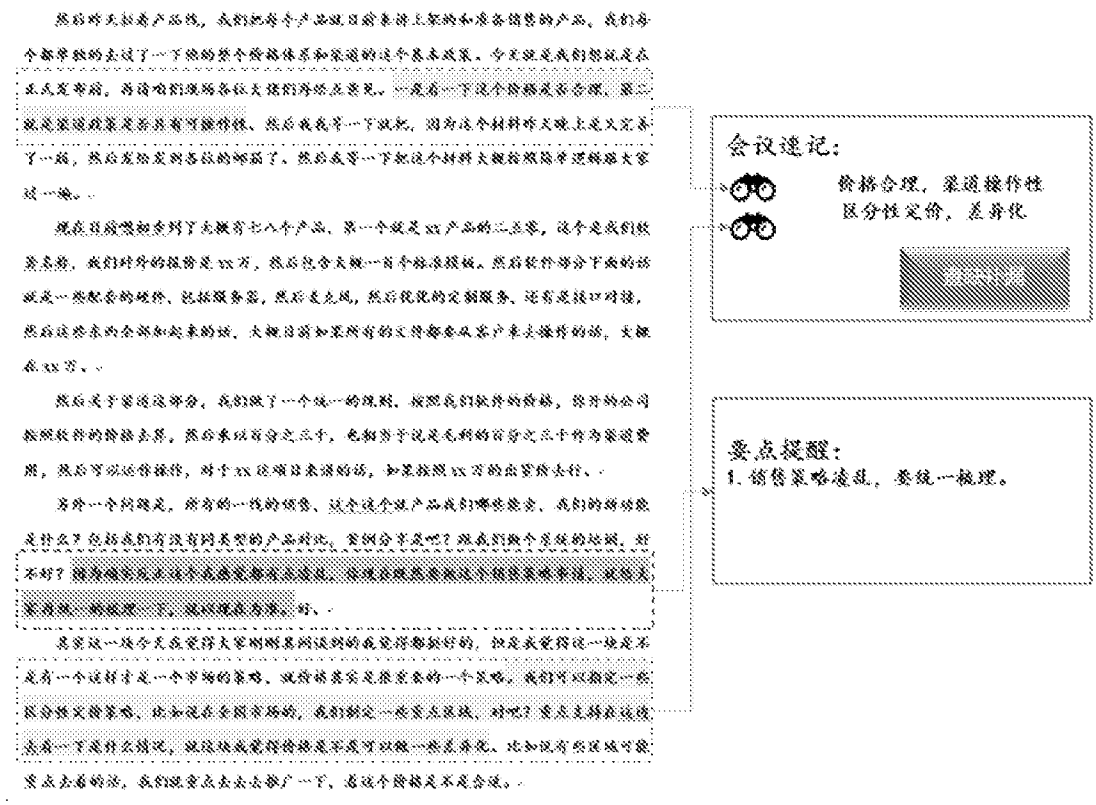


图 2

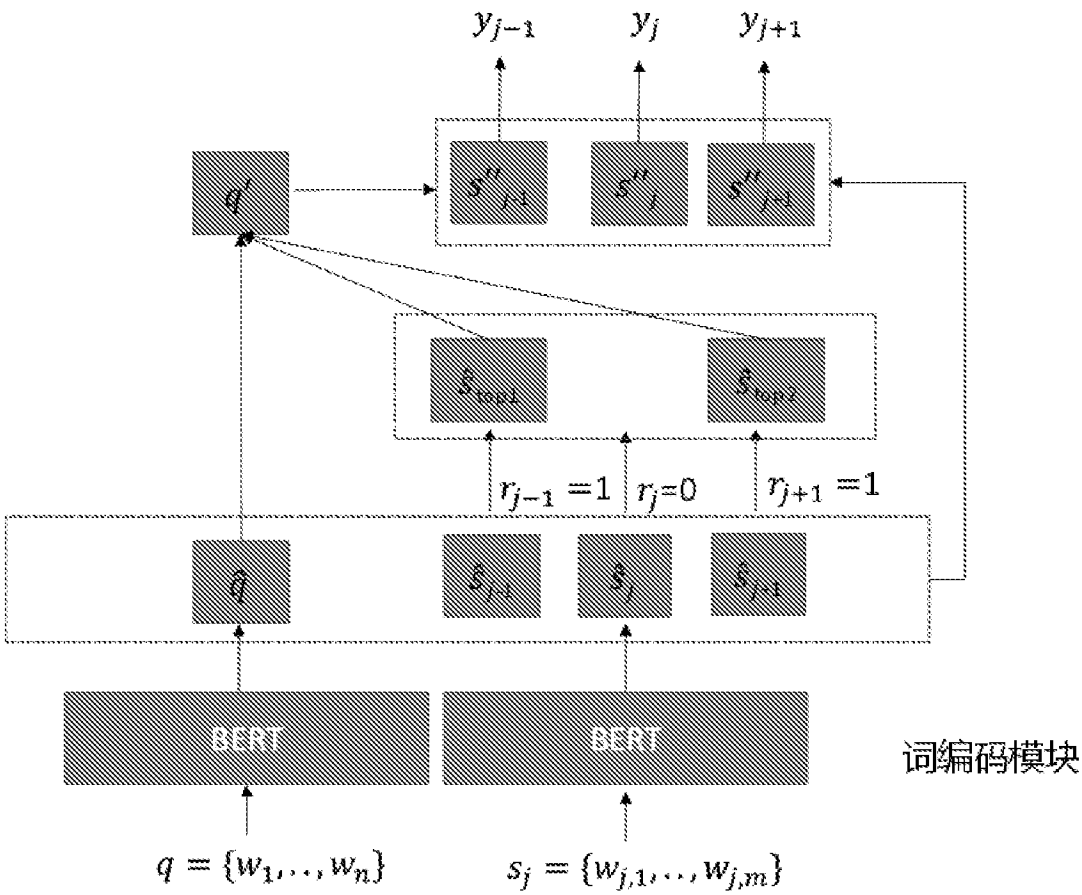

y_{j-1}
y_j
y_{j+1}
q'
s''_{j-1}
s''_j
s''_{j+1}
$s_{top,1}$
$s_{top,2}$
$r_{j-1}=1$
$r_j=0$
$r_{j+1}=1$
q
s_{j-1}
s_j
s_{j+1}
BERT
BERT
词编码模块
$q=\{w_1,..,w_n\}$
$s_j=\{w_{j,1},..,w_{j,m}\}$

图 3

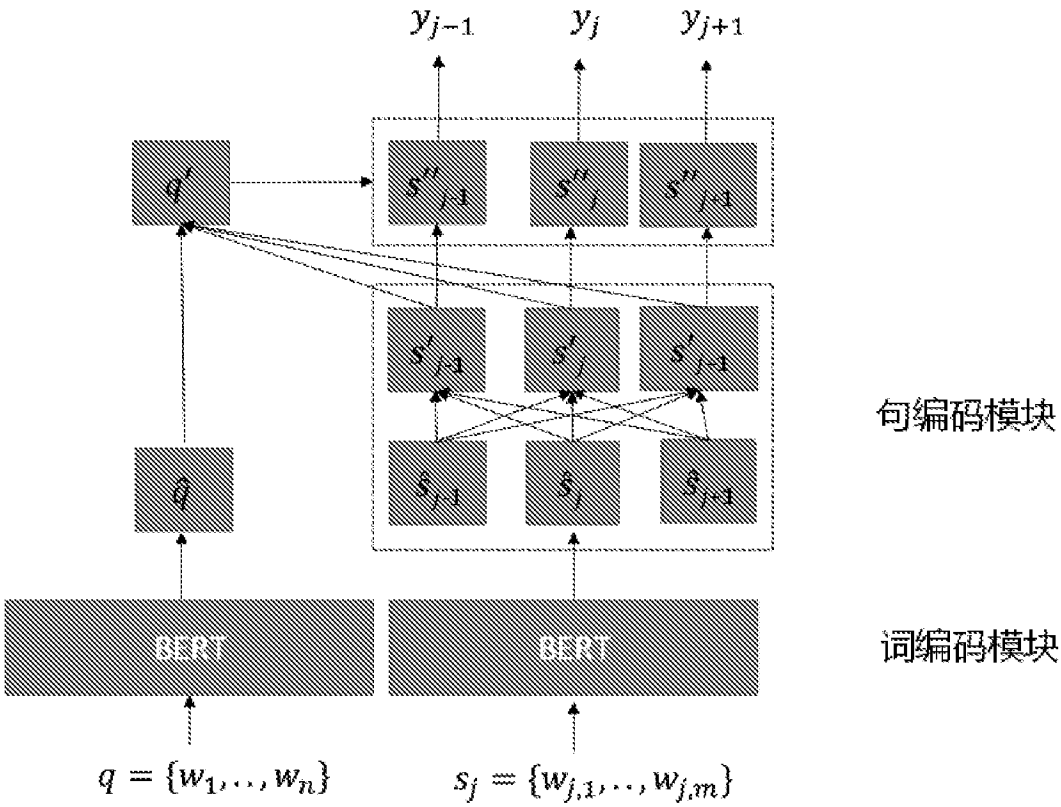

y_{j-1}
y_j
y_{j+1}
q'
s''_{j-1}
s''_j
s''_{j+1}
s'_{j-1}
s'_j
s'_{j+1}
$\hat{q}$
$\hat{s}_{j-1}$
$\hat{s}_j$
$\hat{s}_{j+1}$
句编码模块
BERT
BERT
词编码模块
$q = \{w_1, .., w_n\}$
$s_j = \{w_{j,1}, .., w_{j,m}\}$

图 4

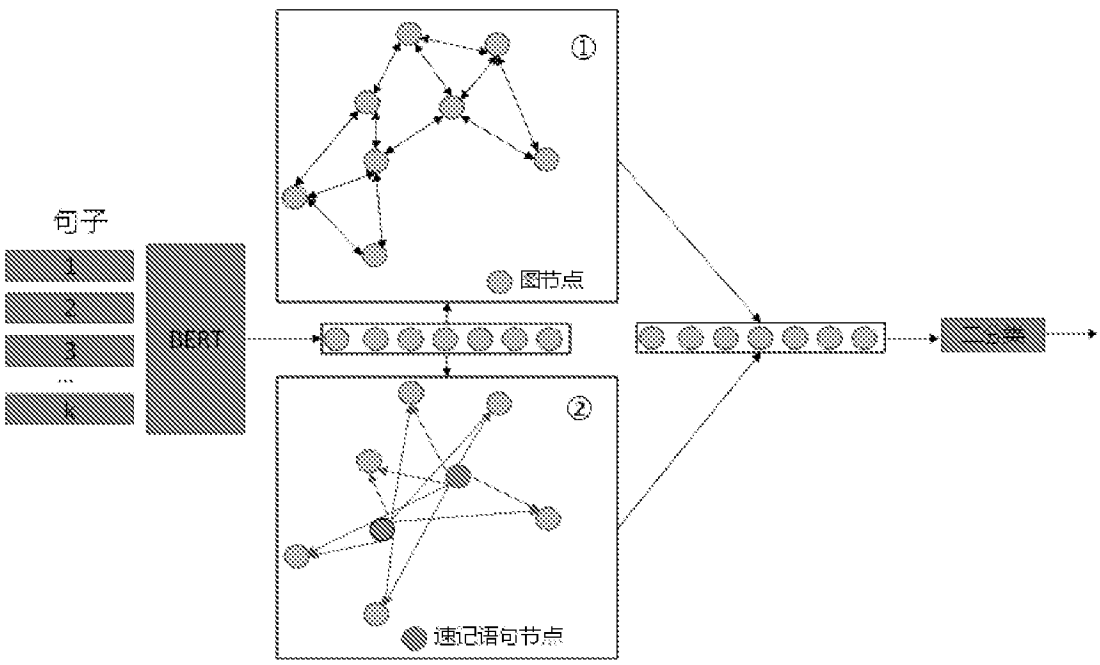

句子
1
2
3
...
k
BERT
①
图节点
②
速记语句节点

图 5

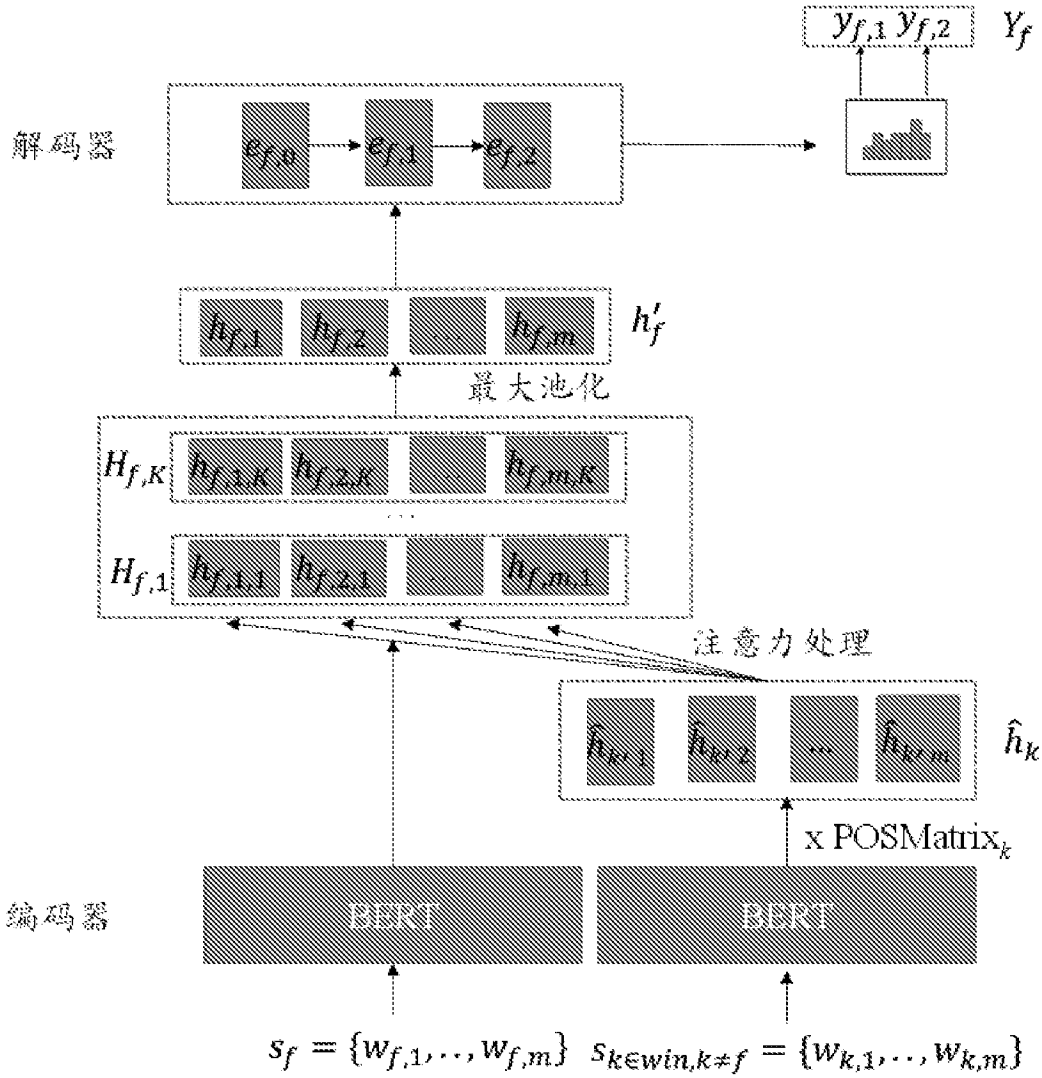

$y_{f,1}$ $y_{f,2}$
Y_f
解码器
$e_{f,0}$
$e_{f,1}$
$e_{f,2}$
$h_{f,1}$ $h_{f,2}$ ... $h_{f,m}$
h'_f
最大池化
$H_{f,K}$
$h_{f,1,K}$ $h_{f,2,K}$ ... $h_{f,m,K}$
$H_{f,1}$
$h_{f,1,1}$ $h_{f,2,1}$ ... $h_{f,m,1}$
注意力处理
$\hat{h}_{k,1}$ $\hat{h}_{k,2}$ ... $\hat{h}_{k,m}$
$\hat{h}_k$
x POSMatrix$_k$
编码器
BERT
BERT
$s_f = \{w_{f,1}, .., w_{f,m}\}$
$s_{k \in win, k \neq f} = \{w_{k,1}, .., w_{k,m}\}$

图 6

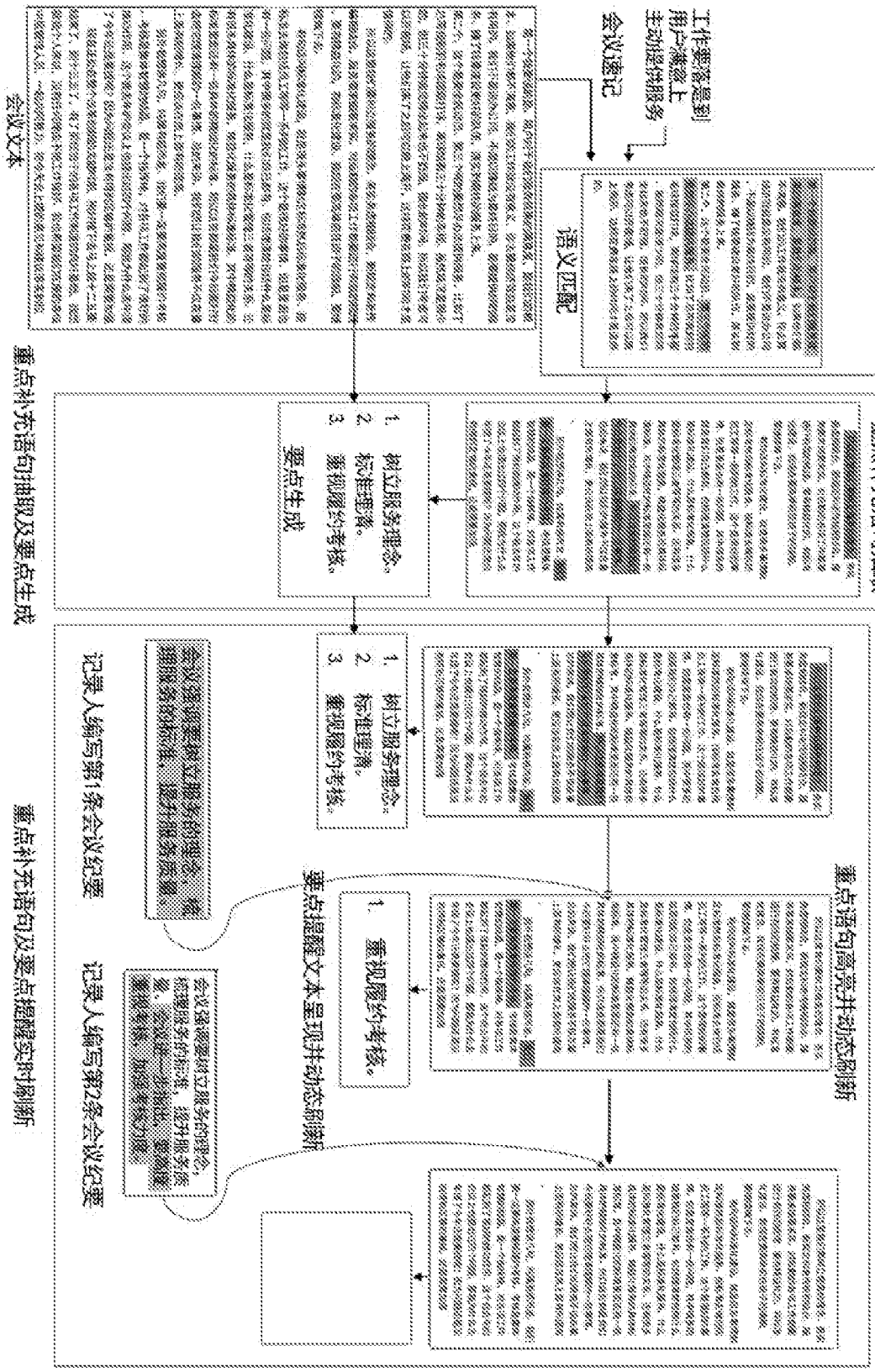

会议文本
工作要落到用户满意上 主动提供服务
会议速记
语义匹配
重点补充语句抽取
重点补充语句抽取及要点生成
1. 树立服务理念。
2. 标准理清。
3. 重视履约考核。
要点生成
重点语句高亮并动态刷新
重点补充语句及要点提醒实时刷新
记录人编写第1条会议纪要
记录人编写第2条会议纪要
要点提醒文本呈现并动态刷新

图 7

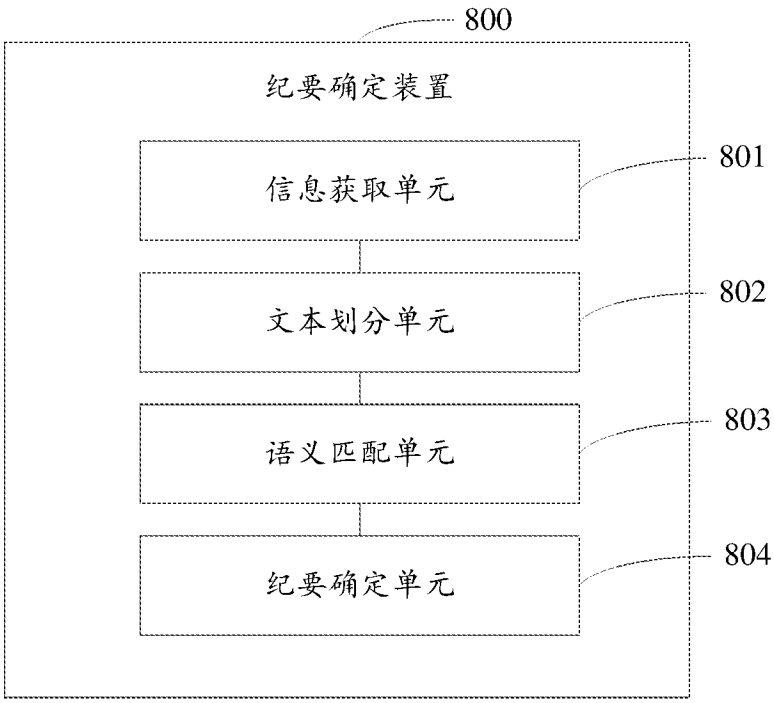


图 8

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2022/133166

A. CLASSIFICATION OF SUBJECT MATTER

G06F16/33(2019.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

VEN, CNABS, CNTXT, WOTXT, USTXT, CNKI, IEEE: 会议, 记录, 纪要, 笔记, 注释, 标注, 文本, 文档, 句, 词, 相似性, 相似度, 语义, 匹配, meeting, minutes, note, text, document, sentence, word, similarity, semantic, match

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 114281948 A (ANHUI TINGJIAN TECHNOLOGY CO., LTD.) 05 April 2022 (2022-04-05) claims 1-30	1-30
PX	CN 114328899 A (IFLYTEK CO., LTD.) 12 April 2022 (2022-04-12) description, paragraphs [0041]-[0232]	1-30
X	CN 109033060 A (IFLYTEK CO., LTD.) 18 December 2018 (2018-12-18) description, paragraphs [0078]-[0199]	1-2, 10, 13-16, 26, 28-30
A	CN 1624686 A (XEROX CORP.) 08 June 2005 (2005-06-08) entire document	1-30
A	US 2013054636 A1 (TANG, D. Y.) 28 February 2013 (2013-02-28) entire document	1-30

☐ Further documents are listed in the continuation of Box C. ☑ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"D" document cited by the applicant in the international application
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
02 February 2023	**10 February 2023**

Name and mailing address of the ISA/CN	Authorized officer
China National Intellectual Property Administration (ISA/CN) **China No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088**	
Facsimile No. **(86-10)62019451**	Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2022/133166

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	114281948	A	05 April 2022	None			
CN	114328899	A	12 April 2022	None			
CN	109033060	A	18 December 2018	None			
CN	1624686	A	08 June 2005	JP	2005174325	A	30 June 2005
				JP	4383328	B2	16 December 2009
				US	2005125219	A1	09 June 2005
				US	7383171	B2	03 June 2008
US	2013054636	A1	28 February 2013	US	8935265	B2	13 January 2015

国际检索报告

国际申请号
PCT/CN2022/133166

A. 主题的分类

G06F16/33(2019.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

IPC: G06F

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称，和使用的检索词（如使用））

VEN, CNABS, CNTXT, WOTXT, USTXT, CNKI, IEEE: 会议，记录，纪要，笔记，注释，标注，文本，文档，句，词，相似性，相似度，语义，匹配，meeting, minutes, note, text, document, sentence, word, similarity, semantic, match

C. 相关文件

类 型*	引用文件，必要时，指明相关段落	相关的权利要求
PX	CN 114281948 A (安徽听见科技有限公司) 2022年4月5日 (2022-04-05) 权利要求1-30	1-30
PX	CN 114328899 A (科大讯飞股份有限公司) 2022年4月12日 (2022-04-12) 说明书第[0041]-[0232]段	1-30
X	CN 109033060 A (科大讯飞股份有限公司) 2018年12月18日 (2018-12-18) 说明书第[0078]-[0199]段	1-2, 10, 13-16, 26, 28-30
A	CN 1624686 A (施乐公司) 2005年6月8日 (2005-06-08) 全文	1-30
A	US 2013054636 A1 (TANG, Ding-yuan) 2013年2月28日 (2013-02-28) 全文	1-30

☐ 其余文件在C栏的续页中列出。

☑ 见同族专利附件。

* 引用文件的具体类型：

“A” 认为不特别相关的表示了现有技术一般状态的文件

“D” 申请人在国际申请中引证的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件，或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件（如具体说明的）

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布，与申请不相抵触，但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件，单独考虑该文件，认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件，当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时，要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期	国际检索报告邮寄日期
2023年2月2日	2023年2月10日

ISA/CN的名称和邮寄地址	受权官员
中国国家知识产权局 中国北京市海淀区蓟门桥西土城路6号 100088	李娜
传真号 (86-10)62019451	电话号码 (+86) 010-53961403

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2022/133166

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	114281948	A	2022年4月5日	无			
CN	114328899	A	2022年4月12日	无			
CN	109033060	A	2018年12月18日	无			
CN	1624686	A	2005年6月8日	JP	2005174325	A	2005年6月30日
				JP	4383328	B2	2009年12月16日
				US	2005125219	A1	2005年6月9日
				US	7383171	B2	2008年6月3日
US	2013054636	A1	2013年2月28日	US	8935265	B2	2015年1月13日