

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

G06F 17/30 (2006.01)

G06F 7/02 (2006.01)



[12] 发明专利申请公布说明书

[21] 申请号 200480036950.4

[43] 公开日 2007年1月3日

[11] 公开号 CN 1890669A

[22] 申请日 2004.10.15

[21] 申请号 200480036950.4

[30] 优先权

[32] 2003.10.15 [33] US [31] 10/687,218

[86] 国际申请 PCT/CA2004/001832 2004.10.15

[87] 国际公布 WO2005/036414 英 2005.4.21

[85] 进入国家阶段日期 2006.6.12

[71] 申请人 施克莱无线公司

地址 加拿大不列颠哥伦比亚

[72] 发明人 约瑟夫·彼得·罗伯特·托赛

[74] 专利代理机构 北京康信知识产权代理有限公司

代理人 余刚

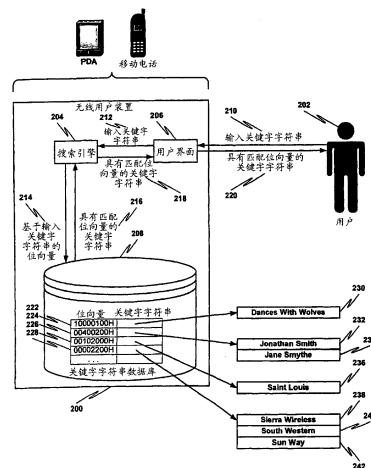
权利要求书 45 页 说明书 22 页 附图 25 页

[54] 发明名称

关键字字符串的增量搜索

[57] 摘要

通过接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串、至少部分地基于输入关键字字符串创建位向量、将位向量与表示至少一个候选关键字字符串的一个或多个其它位向量进行比较以创建一组匹配位向量、并且将常规关键字匹配算法应用于由匹配位向量表示的至少一个候选关键字字符串，可以执行关键字增量搜索。任何匹配候选关键字字符串可以被呈现。



1. 一种用于创建关键字字符串数据库的方法，所述方法包括：

确定一个或多个候选关键字字符串以存储在所述数据库中；

至少部分地基于所述一个或多个候选关键字字符串创建一个或多个位向量，所述一个或多个位向量用于比较输入位向量和所述一个或多个位向量以指示由所述输入位向量代表的输入关键字字符串是否与所述一个或多个候选关键字字符串匹配；以及

将所述一个或多个位向量和指向所述一个或多个候选关键字字符串的引用存储在所述数据库中。
2. 根据权利要求1所述的方法，其中所述位向量还包括至少一个代表非字母数字符号的位。
3. 根据权利要求2所述的方法，其中所述非字母数字符号表示电子邮件地址。
4. 根据权利要求2所述的方法，其中所述非字母数字符号表示移动电话号码。
5. 根据权利要求2所述的方法，其中所述非字母数字符号表示有线电话号码。
6. 根据权利要求2所述的方法，其中所述非字母数字符号表示纸件邮件地址。

7. 根据权利要求2所述的方法,其中所述非字母数字符号表示价格等级。
8. 根据权利要求2所述的方法,其中所述非字母数字符号表示质量等级。
9. 根据权利要求2所述的方法,其中所述非字母数字符号表示烹调风格。
10. 一种用于关键字增量搜索的方法,所述方法包括:
 - 提交包括一个或多个包括一个或多个符号的单词的输入关键字字符串; 以及
 - 响应于所述提交,接收具有与所述输入关键字字符串的位向量匹配的位向量的至少一个候选关键字字符串。
11. 根据权利要求10所述的方法,还包括在预定时间之后取代所述方法。
12. 根据权利要求11所述的方法,其中所述预定时间为两秒钟。
13. 一种用于关键字增量搜索的方法,所述方法包括:
 - 接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串;
 - 至少部分地基于所述输入关键字字符串创建位向量;
 - 将所述位向量与代表至少一个候选关键字字符串的一个或多个其它位向量进行比较,以创建匹配位向量组;
 - 将常规关键字匹配算法应用于由所述匹配位向量组代表的所述至少一个候选关键字字符串; 以及
 - 呈现任何匹配候选关键字字符串。

14. 根据权利要求 13 所述的方法，还包括在预定时间之后取代所述方法。

15. 根据权利要求 14 所述的方法，其中所述预定时间为两秒钟。

16. 根据权利要求 13 所述的方法，其中所述比较与关键字字符串中的关键字词首的顺序无关。

17. 一种用于创建关键字字符串数据库的方法，所述方法包括：

确定一个或多个候选关键字字符串以存储在所述数据库中；

至少部分地基于所述一个或多个候选关键字字符串创建一个或多个位向量，所述位向量具有用于字母表中每个符号的位位置以及对应于代表所述一个或多个候选关键字字符串中单词的第一符号的至少一个符号的位位置的位被置位，所述一个或多个位向量用于比较输入位向量和所述一个或多个位向量，以指示由所述输入位向量代表的输入关键字字符串是否与所述一个或多个候选关键字字符串匹配；以及

将所述一个或多个位向量和指向所述一个或多个候选关键字字符串的引用存储在所述数据库中。

18. 一种用于关键字增量搜索的方法，所述方法包括：

提交包括一个或多个包括一个或多个符号的单词的输入关键字字符串；以及

响应于所述提交，接收至少一个候选关键字字符串，其中在每个所述至少一个候选关键字字符串中至少一个单词的第一符号匹配于所述输入关键字字符串中对应单词的第一符号。

19. 根据权利要求 18 所述的方法，还包括在预定时间之后取代所述方法。

20. 根据权利要求 19 所述的方法，其中所述预定时间为两秒钟。

21. 一种用于关键字增量搜索的方法，所述方法包括：

接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串；

至少部分地基于所述输入关键字字符串创建位向量，所述位向量具有用于字母表中每个符号的位位置和对应于代表所述输入关键字字符串中单词的第一符号的至少一个符号的位置的位被置位；

将所述位向量与代表至少一个候选关键字字符串的一个或多个其它位向量进行比较，以创建匹配位向量组；

将常规关键字匹配算法应用于由所述匹配位向量组代表的所述至少一个候选关键字字符串；以及

呈现任何匹配候选关键字字符串。

22. 根据权利要求 21 所述的方法，还包括在预定时间之后取代所述方法。

23. 根据权利要求 22 所述的方法，其中所述预定时间为两秒钟。

24. 根据权利要求 21 所述的方法，其中所述比较与关键字字符串中的关键字词首的顺序无关。

25. 一种用于比较关键字字符串的方法，所述方法包括：
- 确定一种语言中至少一个符号的相对使用频率；
 - 至少部分地基于所述至少一个字符的相对使用频率，向所述至少一个符号分配统计权重；
 - 将每个所述至少一个符号分配给多个组中的一个；以及
 - 至少部分地基于第一关键字字符串的至少一个符号与第二关键字字符串的至少一个对应符号是否被分配到相同的组，来比较所述第一关键字字符串和所述第二关键字字符串。
26. 根据权利要求 25 所述的方法，其中所述分配还包括将每个所述至少一个符号分配到多个组中的一个，使得用于包括所述多个组中的每个组的符号的统计权重和之间的差别最小。
27. 根据权利要求 25 所述的方法，其中所述相对使用频率包括所述语言中的单词的第一字符中的符号的相对使用频率。
28. 一种用于创建关键字字符串数据库的方法，所述方法包括：
- 确定一个或多个候选关键字字符串以存储在所述数据库中；
 - 至少部分地基于所述一个或多个候选关键字字符串创建一个或多个位向量，所述一个或多个位向量的每位对应于字母表中的一个或多个符号，具有对应于所述一个或多个候选关键字字符串中的单词的第一符号的位位置的位被置位，所述一个或多个位向量用于比较输入位向量和所述一个或多个位向量，以指示由所述输入位向量代表的输入关键字字符串是否与所述一个或多个候选关键字字符串匹配；以及
 - 将所述一个或多个位向量和指向所述一个或多个候选关键字字符串的引用存储在所述数据库中。

29. 一种用于关键字增量搜索的方法，所述方法包括：

提交包括一个或多个包括一个或多个符号的单词的输入关键字字符串，每个符号代表在搜索字符串中的单词的第一符号；以及

响应于所述提交，接收至少一个候选关键字字符串，其中每个候选关键字字符串中的每个单词的第一符号由包括所述一个或多个符号的组构成。

30. 根据权利要求 29 所述的方法，还包括在预定时间之后取代所述方法。

31. 根据权利要求 30 所述的方法，其中所述预定时间为两秒钟。

32. 一种用于关键字增量搜索的方法，所述方法包括：

接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串，每个符号代表搜索字符串中的单词的第一符号；

至少部分地基于所述输入关键字字符串创建位向量，每位对应于字母表中的一个或多个符号，具有对应于所述一个或多个符号的位位置的位被置位；

将所述位向量与代表至少一个候选关键字字符串的一个或多个其它位向量进行比较，以创建匹配位向量组；

将常规关键字匹配算法应用于由所述匹配位向量组代表的所述至少一个候选关键字字符串；以及

呈现任何匹配候选关键字字符串。

33. 根据权利要求 32 所述的方法，还包括在预定时间之后取代所述方法。

34. 根据权利要求 33 所述的方法，其中所述预定时间为两秒钟。
35. 根据权利要求 32 所述的方法，其中所述比较与关键字字符串中的关键字词首的顺序无关。
36. 一种用于创建关键字字符串数据库的方法，所述方法包括：
确定一个或多个候选关键字字符串以存储在所述数据库中；
至少部分地基于所述一个或多个候选关键字字符串创建一个或多个位向量，所述一个或多个位向量的每位对应于字母表中的一个或多个符号，具有对应于所述一个或多个候选关键字字符串中的单词词首的符号的位位置的位被置位，所述一个或多个位向量用于比较输入位向量和所述一个或多个位向量，以指示由所述输入位向量代表的输入关键字字符串是否与所述一个或多个候选关键字字符串匹配；以及
将所述一个或多个位向量和指向所述一个或多个候选关键字字符串的引用存储在所述数据库中。
37. 一种用于关键字增量搜索的方法，所述方法包括：
提交包括一个或多个包括一个或多个符号的单词的输入关键字字符串；以及
响应于所述提交，接收至少一个候选关键字字符串，其中匹配候选关键字字符串的单词词首包括至少一个符号，该符号与所述输入关键字字符串中对应单词的对应符号属于相同符号组。
38. 根据权利要求 37 所述的方法，还包括在预定时间之后取代所述方法。

39. 根据权利要求 38 所述的方法，其中所述预定时间为两秒钟。

40. 一种用于关键字增量搜索的方法，所述方法包括：

接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串；

至少部分地基于所述输入关键字字符串创建位向量，每位对应于字母表中的一个或多个符号，具有对应于所述一个或多个符号中的单词词首的位位置的位被置位；

将所述位向量与代表至少一个候选关键字字符串的一个或多个其它位向量进行比较，以创建匹配位向量组；

将常规关键字匹配算法应用于由所述匹配位向量组代表的所述至少一个候选关键字字符串；以及

呈现任何匹配候选关键字字符串。

41. 根据权利要求 40 所述的方法，还包括在预定时间之后取代所述方法。

42. 根据权利要求 41 所述的方法，其中所述预定时间为两秒钟。

43. 根据权利要求 40 所述的方法，其中所述比较与关键字字符串中的关键字词首的顺序无关。

44. 一种用于关键字增量搜索的方法，所述方法包括：

接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串，每个符号代表搜索字符串中的单词的第一符号；

接收层次，所述层次的元素包括中间节点和叶节点，其代表一个或多个包括一个或多个包括一个或多个符号的单词的关键字字符串；

创建对应于所述层次中的所述一个或多个关键字字符串的层次位向量;

为与所述输入关键字字符串匹配, 搜索所述层次位向量, 所述搜索包括, 为所述层次的每个所述元素:

保存所述输入关键字字符串;

对元素的位向量和至少部分地基于所述输入关键字字符串的位向量应用逻辑“和”操作, 所述应用产生一个结果;

如果所述结果为非零, 从所述输入关键字字符串中除去作为所述元素中单词词首的所述输入关键字字符串中的任何单词;

如果所述输入关键字字符串为空, 将所述元素添加到匹配项列表; 以及

恢复所述输入关键字字符串; 以及

呈报所述匹配项列表。

45. 一种用于创建关键字字符串数据库的方法, 所述方法包括:

确定一个或多个候选关键字字符串以存储在所述数据库中的步骤;

至少部分地基于所述一个或多个候选关键字字符串创建一个或多个位向量的步骤, 所述一个或多个位向量用于比较输入位向量和所述一个或多个位向量以指示由所述输入位向量代表的输入关键字字符串是否与所述一个或多个候选关键字字符串匹配; 以及

将所述一个或多个位向量和指向所述一个或多个候选关键字字符串的引用存储在所述数据库中的步骤。

46. 根据权利要求 45 所述的方法，其中所述位向量还包括至少一个代表非字母数字符号的位。
47. 根据权利要求 46 所述的方法，其中所述非字母数字符号表示电子邮件地址。
48. 根据权利要求 46 所述的方法，其中所述非字母数字符号表示移动电话号码。
49. 根据权利要求 46 所述的方法，其中所述非字母数字符号表示有线电话号码。
50. 根据权利要求 46 所述的方法，其中所述非字母数字符号表示纸件邮件地址。
51. 根据权利要求 46 所述的方法，其中所述非字母数字符号表示价格等级。
52. 根据权利要求 46 所述的方法，其中所述非字母数字符号表示质量等级。
53. 根据权利要求 46 所述的方法，其中所述非字母数字符号表示烹调风格。
54. 一种用于关键字增量搜索的方法，所述方法包括：

提交包括一个或多个包括一个或多个符号的单词的输入关键字字符串的步骤；以及

响应于所述提交，接收具有与所述输入关键字字符串的位向量匹配的位向量的至少一个候选关键字字符串的步骤。

55. 根据权利要求 54 所述的方法，还包括在预定时间之后取代所述方法的步骤。
56. 根据权利要求 55 所述的方法，其中所述预定时间为两秒钟。
57. 一种用于关键字增量搜索的方法，所述方法包括：
接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串的步骤；
至少部分地基于所述输入关键字字符串创建位向量的步骤；
将所述位向量与代表至少一个候选关键字字符串的一个或多个其它位向量进行比较，以创建匹配位向量组的步骤；
将常规关键字匹配算法应用于由所述匹配位向量组代表的所述至少一个候选关键字字符串的步骤；以及
呈现任何匹配候选关键字字符串的步骤。
58. 根据权利要求 57 所述的方法，还包括在预定时间之后取代所述方法的步骤。
59. 根据权利要求 58 所述的方法，其中所述预定时间为两秒钟。
60. 根据权利要求 57 所述的方法，其中所述比较与关键字字符串中的关键字词首的顺序无关。
61. 一种用于创建关键字字符串数据库的方法，所述方法包括：
确定一个或多个候选关键字字符串以存储在所述数据库中的步骤；
至少部分地基于所述一个或多个候选关键字字符串创建一个或多个位向量的步骤，所述位向量具有用于字母表中每个

符号的位位置以及对应于代表所述一个或多个候选关键字字符串中单词的第一符号的至少一个符号的位位置的位被置位,所述一个或多个位向量用于比较输入位向量和所述一个或多个位向量,以指示由所述输入位向量代表的输入关键字字符串是否与所述一个或多个候选关键字字符串匹配; 以及

将所述一个或多个位向量和指向所述一个或多个候选关键字字符串的引用存储在所述数据库中的步骤。

62. 一种用于关键字增量搜索的方法, 所述方法包括:

提交包括一个或多个包括一个或多个符号的单词的输入关键字字符串的步骤; 以及

响应于所述提交, 接收至少一个候选关键字字符串的步骤, 其中在每个所述至少一个候选关键字字符串中至少一个单词的第一符号匹配于所述输入关键字字符串中对应单词的第一符号。

63. 根据权利要求 62 所述的方法, 还包括在预定时间后取代所述方法的步骤。

64. 根据权利要求 63 所述的方法, 其中所述预定时间为两秒钟。

65. 一种用于关键字增量搜索的方法, 所述方法包括:

接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串的步骤;

至少部分地基于所述输入关键字字符串创建位向量的步骤, 所述位向量具有用于字母表中每个符号的位位置和对应于代表所述输入关键字字符串中单词的第一符号的至少一个符号的位置的位被置位;

将所述位向量与代表至少一个候选关键字字符串的一个或多个其它位向量进行比较，以创建匹配位向量组的步骤；

将常规关键字匹配算法应用于由所述匹配位向量组代表的所述至少一个候选关键字字符串的步骤；以及

呈现任何匹配候选关键字字符串的步骤。

66. 根据权利要求 65 所述的方法，还包括在预定时间之后取代所述方法的步骤。

67. 根据权利要求 66 所述的方法，其中所述预定时间量为两秒钟。

68. 根据权利要求 65 所述的方法，其中所述比较与关键字字符串中的关键字词首的顺序无关。

69. 一种用于比较关键字字符串的方法，所述方法包括：

确定一种语言中至少一个符号的相对使用频率的步骤；

至少部分地基于所述至少一个字符的相对使用频率，向所述至少一个符号分配统计权重的步骤；

将每个所述至少一个符号分配给多个组中的一个的步骤；以及

至少部分地基于第一关键字字符串的至少一个符号与第二关键字字符串的至少一个对应符号是否被分配到相同的组，来比较所述第一关键字字符串和所述第二关键字字符串的步骤。

70. 根据权利要求 69 所述的方法，其中所述分配的步骤还包括将每个所述至少一个符号分配到多个组中的一个，使得用于包括所述多个组中的每个组的符号的统计权重和之间的差别最小。

71. 根据权利要求 69 所述的方法, 其中所述相对使用频率包括所述语言中单词的第一字符中的符号的相对使用频率。

72. 一种用于创建关键字字符串数据库的方法, 所述方法包括:

确定一个或多个候选关键字字符串以存储在所述数据库中的步骤;

至少部分地基于所述一个或多个候选关键字字符串创建一个或多个位向量的步骤, 所述一个或多个位向量的每位对应于字母表中的一个或多个符号, 具有对应于所述一个或多个候选关键字字符串中的单词的第一符号的位位置的位被置位, 所述一个或多个位向量用于比较输入位向量和所述一个或多个位向量, 以指示由所述输入位向量代表的输入关键字字符串是否与所述一个或多个候选关键字字符串匹配; 以及

将所述一个或多个位向量和指向所述一个或多个候选关键字字符串的引用存储在所述数据库中。

73. 一种用于关键字增量搜索的方法, 所述方法包括:

提交包括一个或多个包括一个或多个符号的单词的输入关键字字符串, 每个符号代表在搜索字符串中的单词的第一符号的步骤; 以及

响应于所述提交, 接收至少一个候选关键字字符串的步骤, 其中每个候选关键字字符串中的每个单词的第一符号由包括所述一个或多个符号的组构成。

74. 根据权利要求 73 所述的方法, 还包括在预定时间之后取代所述方法的步骤。

75. 根据权利要求 74 所述的方法, 其中所述预定时间为两秒钟。

76. 一种用于关键字增量搜索的方法，所述方法包括：

接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串的步骤，每个符号代表搜索字符串中的单词的第一符号；

至少部分地基于所述输入关键字字符串创建位向量的步骤，每位对应于字母表中的一个或多个符号，具有对应于所述一个或多个符号的位位置的位被置位；

将所述位向量与代表至少一个候选关键字字符串的一个或多个其它位向量进行比较，以创建匹配位向量组的步骤；

将常规关键字匹配算法应用于由所述匹配位向量组代表的所述至少一个候选关键字字符串的步骤；以及

呈现任何匹配候选关键字字符串的步骤。

77. 根据权利要求 76 所述的方法，还包括在预定时间之后取代所述方法的步骤。

78. 根据权利要求 77 所述的方法，其中所述预定时间为两秒钟。

79. 根据权利要求 76 所述的方法，其中所述比较与关键字字符串中的关键字词首的顺序无关。

80. 一种用于创建关键字字符串数据库的方法，所述方法包括：

确定一个或多个候选关键字字符串以存储在所述数据库中的步骤；

至少部分地基于所述一个或多个候选关键字字符串创建一个或多个位向量的步骤，所述一个或多个位向量的每位对应于字母表中的一个或多个符号，具有对应于所述一个或多个候选关键字字符串中的单词词首的符号的位位置的位被置位，所

述一个或多个位向量用于比较输入位向量和所述一个或多个位向量,以指示由所述输入位向量代表的输入关键字字符串是否与所述一个或多个候选关键字字符串匹配; 以及

将所述一个或多个位向量和指向所述一个或多个候选关键字字符串的引用存储在所述数据库中的步骤。

81. 一种用于关键字增量搜索的方法, 所述方法包括:

提交包括一个或多个包括一个或多个符号的单词的输入关键字字符串的步骤; 以及

响应于所述提交, 接收至少一个候选关键字字符串的步骤, 其中匹配候选关键字字符串的单词词首包括至少一个符号, 该符号与所述输入关键字字符串中对应单词的对应符号属于相同符号组。

82. 根据权利要求 81 所述的方法, 还包括在预定时间之后取代所述方法的步骤。

83. 根据权利要求 82 所述的方法, 其中所述预定时间为两秒钟。

84. 一种用于关键字增量搜索的方法, 所述方法包括:

接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串的步骤;

至少部分地基于所述输入关键字字符串创建位向量的步骤, 每位对应于字母表中的一个或多个符号, 具有对应于所述一个或多个符号中的单词词首的位位置的位被置位;

将所述位向量与代表至少一个候选关键字字符串的一个或多个其它位向量进行比较, 以创建匹配位向量组的步骤;

将常规关键字匹配算法应用于由所述匹配位向量组代表的所述至少一个候选关键字字符串的步骤；以及

呈现任何匹配候选关键字字符串的步骤。

85. 根据权利要求 84 所述的方法，还包括在预定时间之后取代所述方法的步骤。

86. 根据权利要求 85 所述的方法，其中所述预定时间为两秒钟。

87. 根据权利要求 84 所述的方法，其中所述比较与关键字字符串中的关键字词首的顺序无关。

88. 一种用于关键字增量搜索的方法，所述方法包括：

接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串的步骤，每个符号代表搜索字符串中的单词的第一符号；

接收层次的步骤，所述层次的元素包括中间节点和叶节点，其代表一个或多个包括一个或多个包括一个或多个符号的单词的关键字字符串；

创建对应于所述层次中的所述一个或多个关键字字符串的层次位向量的步骤；

为与所述输入关键字字符串匹配，搜索所述层次位向量的步骤，所述搜索步骤包括，为所述层次的每个所述元素：

保存所述输入关键字字符串的步骤；

对元素的位向量和至少部分地基于所述输入关键字字符串的位向量应用逻辑“和”操作的步骤，所述应用产生一个结果；

如果所述结果为非零，从所述输入关键字字符串中除去作为所述元素中单词词首的所述输入关键字字符串中的任何单词；

如果所述输入关键字字符串为空，将所述元素添加到匹配项列表的步骤；以及

恢复所述输入关键字字符串的步骤；以及

呈报所述匹配项列表的步骤。

89. 一种机器可读的程序存储装置，包括有可由机器执行的指令程序，以执行一种用于创建关键字字符串数据库的方法，所述方法包括：

确定一个或多个候选关键字字符串以存储在所述数据库中；

至少部分地基于所述一个或多个候选关键字字符串创建一个或多个位向量，所述一个或多个位向量用于比较输入位向量和所述一个或多个位向量以指示由所述输入位向量代表的输入关键字字符串是否与所述一个或多个候选关键字字符串匹配；以及

将所述一个或多个位向量和指向所述一个或多个候选关键字字符串的引用存储在所述数据库中。

90. 根据权利要求 89 所述的程序存储装置，其中所述位向量还包括至少一个代表非字母数字符号的位。
91. 根据权利要求 90 所述的程序存储装置，其中所述非字母数字符号表示电子邮件地址。
92. 根据权利要求 90 所述的程序存储装置，其中所述非字母数字符号表示移动电话号码。

93. 根据权利要求 90 所述的程序存储装置，其中所述非字母数字符号表示有线电话号码。
94. 根据权利要求 90 所述的程序存储装置，其中所述非字母数字符号表示纸件邮件地址。
95. 根据权利要求 90 所述的程序存储装置，其中所述非字母数字符号表示价格等级。
96. 根据权利要求 90 所述的程序存储装置，其中所述非字母数字符号表示质量等级。
97. 根据权利要求 90 所述的程序存储装置，其中所述非字母数字符号表示烹调风格。
98. 一种机器可读的程序存储装置，包括有可由机器执行的指令程序，以执行一种用于关键字增量搜索的方法，所述方法包括：
提交包括一个或多个包括一个或多个符号的单词的输入关键字字符串；以及
响应于所述提交，接收具有与所述输入关键字字符串的位向量匹配的位向量的至少一个候选关键字字符串。
99. 根据权利要求 98 所述的程序存储装置，其中所述方法还包括在预定时间之后取代所述方法。
100. 根据权利要求 99 所述的程序存储装置，其中所述预定时间为两秒钟。

101. 一种机器可读的程序存储装置,包括有可由机器执行的指令程序,以执行一种用于关键字增量搜索的方法,所述方法包括:
- 接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串;
 - 至少部分地基于所述输入关键字字符串创建位向量;
 - 将所述位向量与代表至少一个候选关键字字符串的一个或多个其它位向量进行比较,以创建匹配位向量组;
 - 将常规关键字匹配算法应用于由所述匹配位向量组代表的所述至少一个候选关键字字符串; 以及
 - 呈现任何匹配候选关键字字符串。
102. 根据权利要求 101 所述的程序存储装置,其中所述方法还包括在预定时间之后取代所述方法。
103. 根据权利要求 102 所述的程序存储装置,其中所述预定时间为两秒钟。
104. 根据权利要求 101 所述的程序存储装置,其中所述比较与关键字字符串中的关键字词首的顺序无关。
105. 一种机器可读的程序存储装置,包括有可由机器执行的指令程序,以执行一种用于创建关键字字符串数据库的方法,所述方法包括:
- 确定一个或多个候选关键字字符串以存储在所述数据库中;
 - 至少部分地基于所述一个或多个候选关键字字符串创建一个或多个位向量,所述位向量具有用于字母表中每个符号的位置以及对应于代表所述一个或多个候选关键字字符串中

单词的第一符号的至少一个符号的位位置的位被置位,所述一个或多个位向量用于比较输入位向量和所述一个或多个位向量,以指示由所述输入位向量代表的输入关键字字符串是否与所述一个或多个候选关键字字符串匹配; 以及

将所述一个或多个位向量和指向所述一个或多个候选关键字字符串的引用存储在所述数据库中。

106. 一种机器可读的程序存储装置, 包括有可由机器执行的指令程序, 以执行一种用于关键字增量搜索的方法, 所述方法包括:

提交包括一个或多个包括一个或多个符号的单词的输入关键字字符串; 以及

响应于所述提交, 接收至少一个候选关键字字符串, 其中在每个所述至少一个候选关键字字符串中至少一个单词的第一符号匹配于所述输入关键字字符串中对应单词的第一符号。

107. 根据权利要求 106 所述的程序存储装置, 其中所述方法还包括在预定时间之后取代所述方法。

108. 根据权利要求 107 所述的程序存储装置, 其中所述预定时间为两秒钟。

109. 一种机器可读的程序存储装置, 包括有可由机器执行的指令程序, 以执行一种用于关键字增量搜索的方法, 所述方法包括:

接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串;

至少部分地基于所述输入关键字字符串创建位向量, 所述位向量具有用于字母表中每个符号的位位置和对应于代表

所述输入关键字字符串中单词的第一符号的至少一个符号的位置的位被置位；

将所述位向量与代表至少一个候选关键字字符串的一个或多个其它位向量进行比较，以创建匹配位向量组；

将常规关键字匹配算法应用于由所述匹配位向量组代表的所述至少一个候选关键字字符串；以及

呈现任何匹配候选关键字字符串。

110. 根据权利要求 109 所述的程序存储装置，其中所述方法还包括在预定时间之后取代所述方法。

111. 根据权利要求 110 所述的程序存储装置，其中所述预定时间为两秒钟。

112. 根据权利要求 109 所述的程序存储装置，其中所述比较与关键字字符串中的关键字词首的顺序无关。

113. 一种机器可读的程序存储装置，包括有可由机器执行的指令程序，以执行一种比较关键字字符串的方法，所述方法包括：

确定一种语言中至少一个符号的相对使用频率；

至少部分地基于所述至少一个字符的相对使用频率，向所述至少一个符号分配统计权重；

将每个所述至少一个符号分配给多个组中的一个；以及

至少部分地基于第一关键字字符串的至少一个符号与第二关键字字符串的至少一个对应符号是否被分配到相同的组，来比较所述第一关键字字符串和所述第二关键字字符串。

114. 根据权利要求 113 所述的程序存储装置,其中所述分配还包括将每个所述至少一个符号分配到多个组中的一个,使得用于包括所述多个组中的每个组的符号的统计权重和之间的差别最小。
115. 根据权利要求 113 所述的程序存储装置,其中所述相对使用频率包括所述语言中的单词的第一字符中的符号的相对使用频率。
116. 一种机器可读的程序存储装置,包括有可由机器执行的指令程序,以执行一种用于创建关键字字符串数据库的方法,所述方法包括:
- 确定一个或多个候选关键字字符串以存储在所述数据库中;
- 至少部分地基于所述一个或多个候选关键字字符串创建一个或多个位向量,所述一个或多个位向量的每位对应于字母表中的一个或多个符号,具有对应于所述一个或多个候选关键字字符串中的单词的第一符号的位位置的位被置位,所述一个或多个位向量用于比较输入位向量和所述一个或多个位向量,以指示由所述输入位向量代表的输入关键字字符串是否与所述一个或多个候选关键字字符串匹配; 以及
- 将所述一个或多个位向量和指向所述一个或多个候选关键字字符串的引用存储在所述数据库中。
117. 一种机器可读的程序存储装置,包括有可由机器执行的指令程序,以执行一种用于关键字增量搜索的方法,所述方法包括:
- 提交包括一个或多个包括一个或多个符号的单词的输入关键字字符串,每个符号代表在搜索字符串中的单词的第一符号; 以及

响应于所述提交，接收至少一个候选关键字字符串，其中每个候选关键字字符串中的每个单词的第一符号由包括所述一个或多个符号的组构成。

118. 根据权利要求 117 所述的程序存储装置，其中所述方法还包括在预定时间之后取代所述方法。

119. 根据权利要求 118 所述的程序存储装置，其中所述预定时间为两秒钟。

120. 一种机器可读的程序存储装置，包括有可由机器执行的指令程序，以执行一种用于关键字增量搜索的方法，所述方法包括：

接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串，每个符号代表搜索字符串中的单词的第一符号；

至少部分地基于所述输入关键字字符串创建位向量，每位对应于字母表中的一个或多个符号，具有对应于所述一个或多个符号的位位置的位被置位；

将所述位向量与代表至少一个候选关键字字符串的一个或多个其它位向量进行比较，以创建匹配位向量组；

将常规关键字匹配算法应用于由所述匹配位向量组代表的所述至少一个候选关键字字符串；以及

呈现任何匹配候选关键字字符串。

121. 根据权利要求 120 所述的程序存储装置，其中所述方法还包括在预定时间之后取代所述方法。

122. 根据权利要求 121 所述的程序存储装置，其中所述预定时间为两秒钟。

123. 根据权利要求 120 所述的程序存储装置,其中所述比较与关键字字符串中的关键字词首的顺序无关。

124. 一种机器可读的程序存储装置,包括有可由机器执行的指令程序,以执行一种用于创建关键字字符串数据库的方法,所述方法包括:

确定一个或多个候选关键字字符串以存储在所述数据库中;

至少部分地基于所述一个或多个候选关键字字符串创建一个或多个位向量,所述一个或多个位向量的每位对应于字母表中的一个或多个符号,具有对应于所述一个或多个候选关键字字符串中的单词词首的符号的位位置的位被置位,所述一个或多个位向量用于比较输入位向量和所述一个或多个位向量,以指示由所述输入位向量代表的输入关键字字符串是否与所述一个或多个候选关键字字符串匹配; 以及

将所述一个或多个位向量和指向所述一个或多个候选关键字字符串的引用存储在所述数据库中。

125. 一种机器可读的程序存储装置,包括有可由机器执行的指令程序,以执行一种用于关键字增量搜索的方法,所述方法包括:

提交包括一个或多个包括一个或多个符号的单词的输入关键字字符串; 以及

响应于所述提交,接收至少一个候选关键字字符串,其中匹配候选关键字字符串的单词词首包括至少一个符号,该符号与所述输入关键字字符串中对应单词的对应符号属于相同符号组。

126. 根据权利要求 125 所述的程序存储装置,其中所述方法还包括在预定时间之后取代所述方法。

127. 根据权利要求 126 所述的程序存储装置,其中所述预定时间为两秒钟。
128. 一种机器可读的程序存储装置,包括有可由机器执行的指令程序,以执行一种用于关键字增量搜索的方法,所述方法包括:
接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串;
至少部分地基于所述输入关键字字符串创建位向量,每位对应于字母表中的一个或多个符号,具有对应于所述一个或多个符号中的单词词首的位位置的位被置位;
将所述位向量与代表至少一个候选关键字字符串的一个或多个其它位向量进行比较,以创建匹配位向量组;
将常规关键字匹配算法应用于由所述匹配位向量组代表的所述至少一个候选关键字字符串; 以及
呈现任何匹配候选关键字字符串。
129. 根据权利要求 128 所述的程序存储装置,其中所述方法还包括在预定时间之后取代所述方法。
130. 根据权利要求 129 所述的程序存储装置,其中所述预定时间为两秒钟。
131. 根据权利要求 128 所述的方法,其中所述比较与关键字字符串中的关键字词首的顺序无关。

132. 一种机器可读的程序存储装置,包括有可由机器执行的指令程序,以执行一种用于关键字增量搜索的方法,所述方法包括:

接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串,每个符号代表搜索字符串中的单词的第一符号;

接收层次,所述层次的元素包括中间节点和叶节点,其代表一个或多个包括一个或多个包括一个或多个符号的单词的关键字字符串;

创建对应于所述层次中的所述一个或多个关键字字符串的层次位向量;

为与所述输入关键字字符串匹配,搜索所述层次位向量,所述搜索包括,为所述层次的每个所述元素:

保存所述输入关键字字符串;

对元素位向量和至少部分地基于所述输入关键字字符串的位向量应用逻辑“和”操作,所述应用产生一个结果;

如果所述结果为非零,从所述输入关键字字符串中除去作为所述元素中单词词首的所述输入关键字字符串中的任何单词;

如果所述输入关键字字符串为空,将所述元素添加到匹配项列表;以及

恢复所述输入关键字字符串;以及

呈报所述匹配项列表。

133. 一种用于创建关键字字符串数据库的设备,所述设备包括:

用于确定一个或多个候选关键字字符串以存储在所述数据库中的装置;

用于至少部分地基于所述一个或多个候选关键字字符串创建一个或多个位向量的装置,所述一个或多个位向量用于比较输入位向量和所述一个或多个位向量以指示由所述输入位向量代表的输入关键字字符串是否与所述一个或多个候选关键字字符串匹配; 以及

用于将所述一个或多个位向量和指向所述一个或多个候选关键字字符串的引用存储在所述数据库中的装置。

134. 根据权利要求 133 所述的设备,其中所述位向量还包括至少一个代表非字母数字符号的位。
135. 根据权利要求 134 所述的设备,其中所述非字母数字符号表示电子邮件地址。
136. 根据权利要求 134 所述的设备,其中所述非字母数字符号表示移动电话号码。
137. 根据权利要求 134 所述的设备,其中所述非字母数字符号表示有线电话号码。
138. 根据权利要求 134 所述的设备,其中所述非字母数字符号表示纸件邮件地址。
139. 根据权利要求 134 所述的设备,其中所述非字母数字符号表示价格等级。
140. 根据权利要求 134 所述的设备,其中所述非字母数字符号表示质量等级。
141. 根据权利要求 134 所述的设备,其中所述非字母数字符号表示烹调风格。

142. 一种用于关键字增量搜索的设备，所述设备包括：

用于接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串的装置；

用于至少部分地基于所述输入关键字字符串创建位向量的装置；

用于将所述位向量与代表至少一个候选关键字字符串的一个或多个其它位向量进行比较，以创建匹配位向量组的装置；

用于将常规关键字匹配算法应用于由所述匹配位向量组代表的所述至少一个候选关键字字符串的装置；以及

用于呈现任何匹配候选关键字字符串的装置。

143. 根据权利要求 142 所述的设备，还包括用于在预定时间之后取代所述接收、所述创建、所述比较、所述应用和所述呈现的装置。

144. 根据权利要求 143 所述的设备，其中所述预定时间为两秒钟。

145. 根据权利要求 142 所述的设备，其中所述比较与关键字字符串中的关键字词首的顺序无关。

146. 一种用于创建关键字字符串数据库的设备，所述设备包括：

用于确定一个或多个候选关键字字符串以存储在所述数据库中的装置；

用于至少部分地基于所述一个或多个候选关键字字符串创建一个或多个位向量的装置，所述位向量具有用于字母表中每个符号的位位置以及对应于代表所述一个或多个候选关键字字符串中单词的第一符号的至少一个符号的位位置的位被

置位,所述一个或多个位向量用于比较输入位向量和所述一个或多个位向量,以指示由所述输入位向量代表的输入关键字字符串是否与所述一个或多个候选关键字字符串匹配; 以及

用于将所述一个或多个位向量和指向所述一个或多个候选关键字字符串的引用存储在所述数据库中。

147. 一种用于关键字增量搜索的设备, 所述设备包括:

用于接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串的装置;

用于至少部分地基于所述输入关键字字符串创建位向量的装置, 所述位向量具有用于字母表中每个符号的位位置和对应用于代表所述输入关键字字符串中单词的第一符号的至少一个符号的位置的位被置位;

用于将所述位向量与代表至少一个候选关键字字符串的一个或多个其它位向量进行比较, 以创建匹配位向量组的装置;

用于将常规关键字匹配算法应用于由所述匹配位向量组代表的所述至少一个候选关键字字符串的装置; 以及

用于呈现任何匹配候选关键字字符串的装置。

148. 根据权利要求 147 所述的设备, 还包括用于在预定时间之后, 取代所述接收、所述创建、所述比较、所述应用和所述呈现的装置。

149. 根据权利要求 148 所述的设备, 其中所述预定时间为两秒钟。

150. 根据权利要求 147 所述的设备, 其中所述比较与关键字字符串中的关键字词首的顺序无关。

151. 一种用于比较关键字字符串的设备, 所述设备包括:

用于确定一种语言中至少一个符号的相对使用频率的装置;

用于至少部分地基于所述至少一个字符的相对使用频率, 向所述至少一个符号分配统计权重的装置;

用于将每个所述至少一个符号分配给多个组中的一个的装置; 以及

用于至少部分地基于第一关键字字符串的至少一个符号与第二关键字字符串的至少一个对应符号是否被分配到相同的组, 来比较所述第一关键字字符串和所述第二关键字字符串的装置。

152. 根据权利要求 151 所述的设备, 其中所述用于分配的装置还包括将每个所述至少一个符号分配到多个组中的一个, 使得用于包括所述多个组中的每个组的符号的统计权重和之间的差别最小。

153. 根据权利要求 151 所述的设备, 其中所述相对使用频率包括所述语言中的单词的第一字符中的符号的相对使用频率。

154. 一种用于创建关键字字符串数据库的设备, 所述设备包括:

用于确定一个或多个候选关键字字符串以存储在所述数据库中的装置;

用于至少部分地基于所述一个或多个候选关键字字符串创建一个或多个位向量的装置, 所述一个或多个位向量的每位对应于字母表中的一个或多个符号, 具有对应于所述一个或多个候选关键字字符串中的单词的第一符号的位位置的位被置位, 所述一个或多个位向量用于比较输入位向量和所述一个或

多个位向量,以指示由所述输入位向量代表的输入关键字字符串是否与所述一个或多个候选关键字字符串匹配; 以及

用于将所述一个或多个位向量和指向所述一个或多个候选关键字字符串的引用存储在所述数据库中的装置。

155. 一种用于关键字增量搜索的设备, 所述设备包括:

用于接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串的装置, 每个符号代表搜索字符串中的单词的第一符号;

用于至少部分地基于所述输入关键字字符串创建位向量的装置, 每位对应于字母表中的一个或多个符号, 具有对应于所述一个或多个符号的位位置的位被置位;

用于将所述位向量与代表至少一个候选关键字字符串的一个或多个其它位向量进行比较, 以创建匹配位向量组的装置;

用于将常规关键字匹配算法应用于由所述匹配位向量组代表的所述至少一个候选关键字字符串的装置; 以及

用于呈现任何匹配候选关键字字符串的装置。

156. 根据权利要求 155 所述的设备, 还包括用于在预定时间之后, 取代所述接收、所述创建、所述比较、所述应用和所述呈现的装置。

157. 根据权利要求 156 所述的设备, 其中所述预定时间为两秒钟。

158. 根据权利要求 155 所述的设备, 其中所述比较与关键字字符串中的关键字词首的顺序无关。

159. 一种用于创建关键字字符串数据库的设备，所述设备包括：

用于确定一个或多个候选关键字字符串以存储在所述数据库中的装置；

用于至少部分地基于所述一个或多个候选关键字字符串创建一个或多个位向量的装置，所述一个或多个位向量的每位对应于字母表中的一个或多个符号，具有对应于所述一个或多个候选关键字字符串中的单词词首的符号的位位置的位被置位，所述一个或多个位向量用于比较输入位向量和所述一个或多个位向量，以指示由所述输入位向量代表的输入关键字字符串是否与所述一个或多个候选关键字字符串匹配；以及

用于将所述一个或多个位向量和指向所述一个或多个候选关键字字符串的引用存储在所述数据库中的装置。

160. 一种用于关键字增量搜索的设备，所述设备包括：

用于接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串的装置；

用于至少部分地基于所述输入关键字字符串创建位向量的装置，每位对应于字母表中的一个或多个符号，具有对应于所述一个或多个符号中的单词词首的位位置的位被置位；

用于将所述位向量与代表至少一个候选关键字字符串的一个或多个其它位向量进行比较，以创建匹配位向量组的装置；

用于将常规关键字匹配算法应用于由所述匹配位向量组代表的所述至少一个候选关键字字符串的装置；以及

用于呈现任何匹配候选关键字字符串的装置。

161. 根据权利要求 160 所述的装置, 还包括用于在预定时间之后, 取代所述接收、所述创建、所述比较、所述应用和所述呈现的装置。

162. 根据权利要求 161 所述的设备, 其中所述预定时间为两秒钟。

163. 根据权利要求 160 所述的设备, 其中所述比较与关键字字符串中的关键字词首的顺序无关。

164. 一种用于关键字增量搜索的设备, 所述设备包括:

用于接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串的装置, 每个符号代表搜索字符串中的单词的第一符号;

用于接收层次的装置, 所述层次的元素包括中间节点和叶节点, 其代表一个或多个包括一个或多个包括一个或多个符号的单词的关键字字符串;

用于创建对应于所述层次中的所述一个或多个关键字字符串的层次位向量的装置;

用于为与所述输入关键字字符串匹配, 搜索所述层次位向量的装置, 用于搜索的所述装置包括, 为所述层次的每个所述元素:

用于保存所述输入关键字字符串的装置;

用于对元素的位向量和至少部分地基于所述输入关键字字符串的位向量应用逻辑“和”操作的装置, 所述应用产生一个结果;

如果所述结果为非零, 用于从所述输入关键字字符串中除去作为所述元素中单词词首的所述输入关键字字符串中的任何单词的装置;

如果所述输入关键字字符串为空，用于将所述元素添加到匹配项列表的装置；以及

用于恢复所述输入关键字字符串的装置；以及
用于提供所述匹配项列表的方法。

165. 一种用于创建关键字字符串数据库的装置，所述装置包括：

用于存储所述关键字字符串数据库的存储器；以及
处理器，用于：

确定一个或多个候选关键字字符串以存储在所述数据库中；

至少部分地基于所述一个或多个候选关键字字符串创建一个或多个位向量，所述一个或多个位向量用于比较输入位向量和所述一个或多个位向量，以指示由所述输入位向量代表的输入关键字字符串是否与所述一个或多个候选关键字字符串匹配；以及

将所述一个或多个位向量和指向所述一个或多个候选关键字字符串的引用存储在所述数据库中。

166. 根据权利要求 165 所述的装置，其中所述位向量还包括至少一个代表非字母数字符号的位。

167. 根据权利要求 166 所述的装置，其中所述非字母数字符号表示电子邮件地址。

168. 根据权利要求 166 所述的装置，其中所述非字母数字符号表示移动电话号码。

169. 根据权利要求 166 所述的装置，其中所述非字母数字符号表示有线电话号码。

170. 根据权利要求 166 所述的装置,其中所述非字母数字符号表示纸件邮件地址。

171. 根据权利要求 166 所述的装置,其中所述非字母数字符号表示价格等级。

172. 根据权利要求 166 所述的装置,其中所述非字母数字符号表示质量等级。

173. 根据权利要求 166 所述的装置,其中所述非字母数字符号表示烹调风格。

174. 一种用于关键字增量搜索的装置,所述装置包括:

 用户界面,用于接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串,所述用户界面还用于呈现任何匹配的关键字字符串;以及

 搜索引擎,与所述用户界面通信,并且用于:

 至少部分地基于所述输入关键字字符串创建位向量;

 将所述位向量与代表至少一个候选关键字字符串的一个或多个其它位向量进行比较,以创建匹配位向量组;以及

 将常规关键字匹配算法应用于由所述匹配位向量组代表的所述至少一个候选关键字字符串。

175. 根据权利要求 174 所述的装置,其中所述装置还用于在预定时间之后取代所述搜索引擎。

176. 根据权利要求 175 所述的装置,其中所述预定时间为两秒钟。

177. 根据权利要求 174 所述的装置,其中所述装置还用于比较与关键字字符串中的关键字词首的顺序无关的所述位向量。

178. 一种用于创建关键字字符串数据库的装置,所述装置包括:

存储器,用于存储所述关键字字符串数据库;以及
处理器,用于:

确定一个或多个候选关键字字符串以存储在所述数据库中;

至少部分地基于所述一个或多个候选关键字字符串创建一个或多个位向量,所述位向量具有用于字母表中每个符号的位位置以及对应于代表所述一个或多个候选关键字字符串中单词的第一符号的至少一个符号的位位置的位被置位,所述一个或多个位向量用于比较输入位向量和所述一个或多个位向量,以指示由所述输入位向量代表的输入关键字字符串是否与所述一个或多个候选关键字字符串匹配;以及

将所述一个或多个位向量和指向所述一个或多个候选关键字字符串的引用存储在所述数据库中。

179. 一种用于关键字增量搜索的装置,所述装置包括:

用户界面,用于接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串,所述用户界面还用于呈现任何匹配的关键字字符串;以及

搜索引擎,与所述用户界面通信,并且用于:

至少部分地基于所述输入关键字字符串创建位向量,所述位向量具有用于字母表中每个符号的位位置 and 对应于代表所述输入关键字字符串中单词的第一符号的至少一个符号的位置的位被置位;

将所述位向量与代表至少一个候选关键字字符串的一个或多个其它位向量进行比较，以创建匹配位向量组；以及

将常规关键字匹配算法应用于由所述匹配位向量组代表的所述至少一个候选关键字字符串。

180. 根据权利要求 179 所述的装置，其中所述装置还用于在预定时间之后取代所述搜索引擎。

181. 根据权利要求 180 所述的装置，其中所述预定时间为两秒钟。

182. 根据权利要求 179 所述的装置，其中所述装置还用于比较与关键字字符串中的关键字词首的顺序无关的所述位向量。

183. 一种用于比较关键字字符串的装置，所述装置包括：

存储器，用于存储一种语言中至少一个符号的至少一个相对使用频率；以及

处理器，用于：

确定一种语言中至少一个符号的相对使用频率；

至少部分地基于所述至少一个字符的相对使用频率，向所述至少一个符号分配统计权重；

将每个所述至少一个符号分配给多个组中的一个；以及

至少部分地基于第一关键字字符串的至少一个符号与第二关键字字符串的至少一个对应符号是否被分配到相同的组，来比较所述第一关键字字符串和所述第二关键字字符串。

184. 根据权利要求 183 所述的装置,其中所述处理器还用于将每个所述至少一个符号分配给多个组中的一个,使得用于包括所述多个组中的每个组的符号的统计权重和之间的差别最小。

185. 根据权利要求 183 所述的装置,其中所述相对使用频率包括所述语言中的单词的第一字符中的符号的相对使用频率。

186. 一种用于创建关键字字符串数据库的装置,所述装置包括:

存储器,用于存储所述关键字字符串数据库;以及

处理器,用于:

确定一个或多个候选关键字字符串以存储在所述数据库中;

至少部分地基于所述一个或多个候选关键字字符串创建一个或多个位向量,所述一个或多个位向量的每位对应于字母表中的一个或多个符号,具有对应于所述一个或多个候选关键字字符串中的单词的第一符号的位位置的位被置位,所述一个或多个位向量用于比较输入位向量和所述一个或多个位向量,以指示由所述输入位向量代表的输入关键字字符串是否与所述一个或多个候选关键字字符串匹配;以及

将所述一个或多个位向量和指向所述一个或多个候选关键字字符串的引用存储在所述数据库中。

187. 一种用于关键字增量搜索的装置,所述装置包括:

用户界面,用于接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串,每个符号代表搜索字符串中的单词的第一符号,所述用户界面还用于呈现任何匹配的关键字字符串;以及

搜索引擎，与所述用户界面通信，并且用于：

至少部分地基于所述输入关键字字符串创建位向量，每位对应于字母表中的一个或多个符号，具有对应于所述一个或多个符号的位位置的位被置位；

将所述位向量与代表至少一个候选关键字字符串的一个或多个其它位向量进行比较，以创建匹配位向量组；以及

将常规关键字匹配算法应用于由所述匹配位向量组代表的所述至少一个候选关键字字符串。

188. 根据权利要求 187 所述的装置，其中所述装置还用于在预定时间之后取代所述搜索引擎。

189. 根据权利要求 188 所述的装置，其中所述预定时间为两秒钟。

190. 根据权利要求 187 所述的装置，其中所述装置还用于比较与关键字字符串中的关键字词首的顺序无关的所述位向量。

191. 一种用于创建关键字字符串数据库的装置，所述装置包括：

存储器，用于存储所述关键字字符串数据库；以及
处理器，用于：

确定一个或多个候选关键字字符串以存储在所述数据库中；

至少部分地基于所述一个或多个候选关键字字符串创建一个或多个位向量，所述一个或多个位向量的每位对应于字母表中的一个或多个符号，具有对应于所述一个或多个候选关键字字符串中的单词词首的符号的位位置的位被置位，所述一个或多个位向量用于比较输入位

向量和所述一个或多个位向量，以指示由所述输入位向量代表的输入关键字字符串是否与所述一个或多个候选关键字字符串匹配；以及

将所述一个或多个位向量和指向所述一个或多个候选关键字字符串的引用存储在所述数据库中。

192. 一种用于关键字增量搜索的装置，所述装置包括：

用户界面，用于接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串，所述用户界面还用于呈现任何匹配的关键词字符串；以及

搜索引擎，与所述用户界面通信，并且用于：

至少部分地基于所述输入关键字字符串创建位向量，每位对应于字母表中的一个或多个符号，具有对应于所述一个或多个符号中的单词词首的位位置的位被置位；

将所述位向量与代表至少一个候选关键字字符串的一个或多个其它位向量进行比较，以创建匹配位向量组；以及

将常规关键字匹配算法应用于由所述匹配位向量组代表的所述至少一个候选关键字字符串。

193. 根据权利要求 192 所述的装置，其中所述装置还用于在预定时间之后取代所述搜索引擎。

194. 根据权利要求 193 所述的装置，其中所述预定时间为两秒钟。

195. 根据权利要求 192 所述的装置，其中所述装置还用于比较与关键字字符串中的关键词词首的顺序无关的所述位向量。

196. 一种用于关键字增量搜索的装置，所述装置包括：

用户界面，用于接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串，所述用户界面还用于呈现任何匹配的关键字字符串；以及

搜索引擎，与所述用户界面通信，并且用于：

接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串，每个符号代表搜索字符串中的单词的第一符号；

接收层次，所述层次的元素包括中间节点和叶节点，其代表一个或多个包括一个或多个包括一个或多个符号的单词的关键字字符串；

创建对应于所述层次中的所述一个或多个关键字字符串的层次位向量；

为与所述输入关键字字符串匹配，搜索所述层次位向量，所述装置还用于，为所述层次的每个所述元素：

保存所述输入关键字字符串；

对元素的位向量和至少部分地基于所述输入关键字字符串的位向量应用逻辑“和”操作，所述应用产生一个结果；

如果所述结果为非零，从所述输入关键字字符串中除去作为所述元素中单词词首的所述输入关键字字符串中的任何单词；

如果所述输入关键字字符串为空，将所述元素添加到匹配项列表；以及

恢复所述输入关键字字符串；以及

呈报所述匹配项列表；以及

将常规关键字匹配算法应用于由所述匹配项列表中的一个或多个元素所表示的所述至少一个关键字字符串。

197. 一种用于比较关键字字符串的方法，所述方法包括：

将一种语言中的至少一个符号中的每个分配给多个组中的一个；以及

至少部分地基于第一关键字字符串的至少一个符号与第二关键字字符串的至少一个对应符号是否被分配到相同的组，来比较所述第一关键字字符串和所述第二关键字字符串。

198. 根据权利要求 197 所述的方法，其中所述多个组对应于电话键盘符号分组。

199. 一种用于比较关键字字符串的方法，所述方法包括：

将一种语言中的至少一个符号中的每个分配给多个组中的一个的步骤；以及

至少部分地基于第一关键字字符串的至少一个符号与第二关键字字符串的至少一个对应符号是否被分配到相同的组，来比较所述第一关键字字符串和所述第二关键字字符串的步骤。

200. 根据权利要求 199 所述的方法，其中所述多个组对应于电话键盘符号分组。

201. 一种机器可读的程序存储装置,包括有可由机器执行的指令程序,以执行一种用于比较关键字字符串的方法,所述方法包括:

将一种语言中的至少一个符号中的每个分配给多个组中的一个; 以及

至少部分地基于第一关键字字符串的至少一个符号与第二关键字字符串的至少一个对应符号是否被分配到相同的组,来比较所述第一关键字字符串和所述第二关键字字符串。

202. 根据权利要求 201 所述的程序存储装置,其中所述多个组对应于电话键盘符号分组。

203. 一种用于比较关键字字符串的设备,所述设备包括:

用于将一种语言中的至少一个符号中的每个分配给多个组中的一个的装置; 以及

用于至少部分地基于第一关键字字符串的至少一个符号与第二关键字字符串的至少一个对应符号是否被分配到相同的组,来比较所述第一关键字字符串和所述第二关键字字符串的装置。

204. 根据权利要求 203 所述的装置,其中所述多个组对应于电话键盘符号分组。

205. 一种用于比较关键字字符串的装置,所述装置包括:

存储器,用于存储所述关键字字符串; 以及

处理器,用于:

将一种语言中的至少一个符号中的每个分配给多个组中的一个; 以及

至少部分地基于第一关键字字符串的至少一个符号与第二关键字字符串的至少一个对应符号是否被分配到相同的组，来比较所述第一关键字字符串和所述第二关键字字符串。

206. 根据权利要求 205 所述的装置，其中所述多个组对应于电话键盘符号分组。

关键字字符串的增量搜索

技术领域

本发明涉及计算机科学领域。更具体地，本发明涉及一种用于关键字字符串的增量搜索的方法和装置。

背景技术

通常，无线用户装置，例如个人数字助理（PDA）和移动电话，包括一个小键盘。键盘的尺寸导致使用该键盘很困难。例如，期望找到题为“Dances with Wolves”的电影的PDA用户将不得不输入几个符号来执行有关该电影的信息的搜索。因此，期望找到这样输入的关键字字符串的答案，而不必键入整个搜索字符串。

一些解决方案对用户键入的每个输入的关键字字符串提供实时响应。如果装置可以在每个按键被键入后以匹配列表来响应，并且可以与用户键入一样迅速，则用户将知道何时停止键入按键。作为以该方式工作的系统的实例，假设数据库包括大量的关键字字符串，例如：

“Dances with Wolves”

“Jonathan Smith”

“Saint Louis”

“Sierra Wireless”

当用户在输入窗口中键入“w”时，系统立即以具有至少一个以符号“w”开始的关键字的关键字字符串来响应：

“Dances with Wolves”

“Sierra Wireless”

当用户继续键入“w d”时，系统进一步将响应缩小到具有至少一个以符号“w”开始的关键字以及至少一个以符号“d”开始的关键字的关键字字符串。在以上的实例中，只有一个可能的匹配：

“Dance with Wolves”。

然而，当前的实时解决方案具有降低无线用户装置的有效性的缺陷。一个“实时”解决方案将所搜索的数据库的大小限制在 10,000 条目（例如装置上的联系人）。另一个“实时”解决方案将关键字搜索限制于在关键字字符串中的单个关键字。

其它解决方案提供了对相对较大数据库的穷举搜索能力。然而，这些解决方案不可实时进行，经常需要几分钟去完成。

对用于实时关键字字符串匹配的系统和方法的技术存在要求，该技术使受资源约束和受 CPU 约束的装置（例如移动电话、PDA、或类似物）能够与用户在用户装置上键入按键一样迅速地找到匹配项。对这种解决方案存在的另外一个要求是相对的有效率，对这种解决方案存在的另外一个要求是其允许关键字搜索字符串中多于一个单词。对这种解决方案仍存在的另外一个需要是其可以搜索相对较多数量的条目的能力。

发明内容

通过接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串、基于至少部分输入的关键字字符串创建位向量、将

该位向量与表示至少一个候选关键字字符串的一个或多个其它位向量进行比较以创建一组匹配位向量、并且将常规关键字匹配算法应用于由匹配位向量表示的至少一个候选关键字字符串，可以执行关键字增量搜索。呈现任何匹配候选关键字字符串。

附图说明

结合在说明书中并构成说明书的一部分的附图，示出了本发明的一个或多个实施例，并且与具体的描述一起，用来解释本发明的原理与实施方法。

在图中：

图 1 是适用于本发明的实施的计算机系统的方框图。

图 2 是根据本发明的实施例的用于关键字字符串的增量搜索的无线用户装置的方框图。

图 3 是根据本发明的一个实施例的示出了用于创建关键字字符串数据库的方法的流程图。

图 4 是根据本发明的一个实施例的从用户角度示出了用于关键字字符串的增量搜索的方法的流程图。

图 5 是根据本发明的一个实施例的从搜索引擎角度示出了用于关键字字符串的增量搜索的方法的流程图。

图 6A 是根据本发明的一个实施例的示出了根据输入的字符串“Dances with Wolves”创建位向量的方框图。

图 6B 是根据本发明的一个实施例的示出了根据输入的字符串“Jonathan Smith”创建位向量的方框图。

图 6C 是根据本发明的一个实施例的示出了根据输入的字符串“Saint Louis”创建位向量的方框图。

图 6D 是根据本发明的一个实施例的示出了根据输入的字符串“Sierra Wireless”创建位向量的方框图。

图 7 是根据本发明的一个实施例的示出了用于创建包括位向量的关键字字符串数据库的方法的流程图，该位向量具有用于字母表中每个符号的位位置，以及对应于表示单词的第一符号的至少一个符号的位位置的位被置位。

图 8 是根据本发明的一个实施例的从使用图 7 的数据库的用户的角度示出了用于关键字字符串的增量搜索的方法的流程图。

图 9 是根据本发明的一个实施例的从使用图 7 的数据库的搜索引擎的角度示出了用于关键字字符串的增量搜索的方法的流程图。

图 10 是示出了以使用每个符号的相对频率为基础创建一个或多个符号的统计分组的方框图。

图 11 是示出了根据本发明的一个实施例的基于使用图 10 的统计分组的输入关键字字符串中关键字的第一符号创建位向量的方框图。

图 12 是示出了根据本发明的一个实施例的用于以符号的统计分组中的符号的成员资格为基础比较关键字字符串的方法的流程图。

图 13 是示出了根据本发明的一个实施例的用于创建包括一个或多个候选关键字字符串的总位向量的关键字字符串数据库的方法的流程图。

图 14 是从使用图 13 的数据库的用户的角度示出了根据本发明的一个实施例的用于关键字字符串的增量搜索的方法的流程图。

图 15 是从使用图 13 的数据库的搜索引擎的角度示出了根据本发明的一个实施例的用于关键字字符串的增量搜索的方法的流程图。

图 16 是示出了根据本发明的一个实施例的根据符号统计分组中的关键字字符串的第一单词的第一符号的成员资格创建总位向量的方框图。

图 17 是示出了根据本发明的一个实施例的根据符号统计分组中的关键字字符串的第一单词的第一符号的成员资格创建总位向量的方框图。

图 18 是示出了根据本发明的一个实施例的用于创建关键字字符串数据库的方法的流程图，该数据库包括一个或多个总位向量，其中每位对应于字母表中的一个或多个符号，并且其中具有对应于在候选关键字字符串中的单词的词首中的符号的位位置的位被置位。

图 19 是从使用图 18 的数据库的用户的角度示出了根据本发明的一个实施例的关键字字符串的增量搜索的方法的流程图。

图 20 是从使用图 18 的数据库的搜索引擎的角度示出了根据本发明的一个实施例的关键字字符串的增量搜索的方法的流程图。

图 21 是示出了根据本发明的一个实施例的将总位向量的至少一位分配给与关键字字符串相关的至少一个非字母属性的方框图。

图 22 是示出了根据本发明的一个实施例的将总位向量的至少一位分配给与关键字字符串相关的非字母属性的方框图。

图 23A 是示出了根据本发明的一个实施例的关键字字符串的分层组织的方框图。

图 23B 是示出了根据本发明的一个实施例的关键字字符串的分层组织的方框图。

图 23C 是示出了根据本发明的一个实施例的关键字字符串的分层组织的方框图。

图 24 是示出了根据本发明的一个实施例的分层构成的关键字字符串的增量搜索的方法的流程图。

图 25 是示出了根据本发明的一个实施例的搜索与模式匹配的位向量的层次的方法的流程图。

具体实施方式

在此，本文中的本发明的实施例描述了关键字字符串的增量搜索的方法和装置。那些本领域的普通技术人员将认识到下面对本发明的详细描述仅作为示例，并不用于任何方式的限制。从本公开收益的技术人员将容易地想到本发明的其它实施例。将详细参考在附图中示出的本发明的实施。在附图及下面涉及相同或类似部分的详细描述中始终使用相同的附图标号。

为了清楚的目的，没有示出及描述在此描述的实施方式的全部常规特征。当然，在开发任何这样的实际实施方式中应该理解，为了实现开发者的特定目标，必须做出许多特定实施方式的决定，例如顺应应用和业务相关的限制，并且随着一个实施方式到另一个实

施方式和一个开发者到另一个开发者，这些特定目标将发生变化。而且，可以意识到这样的开发工作可能是复杂及耗时的，但是对那些获益于本公开的本领域的普通技术人员将不过是工程的常规工作。

根据本发明的一个实施例，将使用多种类型的操作系统(OS)、计算平台、固件、计算机程序、计算机语言、和/或通用机器来实现部件、处理步骤、和/或数据结构。方法可以作为运行在处理电路上的被编程的处理来运行。处理电路可以采取处理器和操作系统的多种组合的形式，或独立装置的形式。处理可以由硬件所执行的指令、单独的硬件、或它们的组合来实施。软件可以被存储在机器可读的程序存储装置上。

另外，那些本领域的普通技术人员将认识到可以使用较少通用性质的装置而不背离在此公开的发明概念的范畴和精神，例如硬件化的装置、现场可编程逻辑器件(FPLDs)，包括现场可编程门阵列(FPGAs)和复杂可编程逻辑器件(CPLDs)、专用集成电路(ASICs)、或类似物。

在本发明的上下文中，术语“网络”包括局域网、广域网、互联网、有线电视系统、电话系统、无线通信系统、光纤网络、ATM网络、帧中继网络、卫星通信系统等。在本领域中这些网络都是熟知的并且因此不在此进一步描述。

在本发明的上下文中，术语“无线用户装置”包括任何可以进行无线通信以及可以向用户呈现信息的装置。例如，信息可以通过视觉、听觉、或触觉被呈现。例如，示例性的无线用户装置包括个人数字助理(PDA)、移动电话、个人计算机、寻呼机等。

在本发明的上下文中，术语“字母表”包括用一种或多种语言写成的一组字母或其它符号。

在本发明的上下文中，术语“符号”包括文字数字式字符，例如数字或字母表的字母。更通常地，术语“符号”包括在涉及特定领域的书写或印刷中，表示操作、数量、元素、关系、或性质所使用的标记。

在本发明的上下文中，术语“常规关键字匹配算法”定义为强力匹配算法，其中完整的字符串直接与一个或多个字符串进行比较。

图1描述了适用于本发明的实现的计算机系统**100**的方框图。如图1所示，系统**100**包括互联主要子系统的总线**102**，例如处理器**104**、内部存储器**106**（例如RAM）、输入/输出（I/O）控制器**108**、可移动存储器**122**（例如存储器卡）、例如通过显示适配器**112**的显示屏**110**的外部装置、滚动类型输入装置**114**、游戏杆**116**、数字键盘**118**、字母数字式键盘**118**、定向导航键**126**和无线接口**120**。可以连接许多其它装置。无线网络接口**120**、有线网络接口**128**、或两者，可以使用本领域技术人员所知的网络接口系统连接局域网或广域网（例如互联网）。

许多其它装置或子系统（没有示出）可以以类似的方式被连接。同样，图1中示出的所有装置没有必要都存在以实现本发明。此外，装置和子系统可以以与图1中示出的方式不同的方式互联。实现本发明的代码可以以可操作的方式放置在内部存储器**106**中或存储在诸如可移动存储器**122**、软盘或CD-ROM的存储介质上。

现在转到图2，图2示出了根据本发明的实施例用于关键字字符串的增量搜索的无线用户装置的方框图。无线用户装置**200**包括

用户界面 206，其配置用于接收来自用户 202 的输入的关键字字符串 210 并且向用户 202 提供一个或多个具有匹配位向量 220 的关键字字符串。无线用户装置 200 还包括搜索引擎 204，其用于接收用户界面 206 的输入关键字字符串 212，从关键字字符串数据库 208 获得一个或多个具有与基于输入的关键字字符串 214 的位向量匹配的位向量的关键字字符串。搜索引擎 204 同时也用于将匹配的关键字字符串传送到用户界面 206 用于接下来提供给用户 202。关键字字符串数据库 208 包括一个或多个包括位向量的条目以及与该位向量相关联的一个或多个关键字字符串的引用。如图 2 所示，位向量 222 至 228 包括 32 位数字，其中英语字母表的每个符号都分配了一个特定位位置，并且位向量中被置位的位（具有值“1”）指示在匹配关键字字符串中至少一个单词的第一符号等于与该位位置相关联的符号。以下参考图 6 至 9 更详细地解释了该特定位向量映射和其它示例性位向量映射。

在操作中，用户 202 通过无线用户装置 200 的用户界面 206 输入包括一个或多个包括一个或多个符号的关键字。用户界面 206 接收输入的关键字字符串 210 并且将其传送到搜索引擎 204。搜索引擎 204 确定输入的关键字字符串的位向量并且与关键字字符串数据库 208 通信以接收与匹配位向量相关联的一个或多个关键字字符串。然后，搜索引擎 204 将一个或多个关键字字符串传送到用户界面 206，向用户 202 显示匹配关键字字符串。

根据本发明的一个实施例，无线用户装置 200 被配置为在每次用户 202 输入一个输入的关键字字符串 210 的符号时触发关键字字符串搜索。用户 202 可以继续输入用于搜索的符号，匹配的数量随着每次输入而减少。

许多其它装置或子系统（没有示出）可以以类似于图 2 所示的方式被连接。同样，图 2 中示出的所有装置没有必要都存在以实现

本发明。此外，装置和子系统可以以与图 2 中示出的方式不同的方式被互联。

图 3 至图 5 在高层示出了关键字字符串数据库的创建和使用。图 6A 至图 9 示出了包括位向量的关键字字符串数据库的创建和使用，该位向量具有用于字母表中每个符号的位位置，并且使对应于代表单词的第一符号的至少一个符号的位位置的位被置位。图 10 至图 12 示出了基于在符号的统计分组中的符号的成员资格比较关键字字符串。图 13 至图 15 示出了创建和使用包括候选关键字字符串的一个或多个总位向量的关键字字符串数据库。图 16 至图 20 示出了创建和使用包括一个或多个总位向量的关键字字符串数据库，其中每个位对应于字母表中的一个或多个符号，并且其中具有对应于在候选关键字字符串中的单词的词首中的符号的位位置的位被置位。

现在转到图 3，示出了根据本发明的一个实施例的创建关键字字符串数据库的方法的流程图。在 **300**，确定一个或多个候选关键字字符串以存储在数据库中。候选关键字字符串包括与所有来自用户的输入关键字字符串可匹配的关键字字符串。为每个在 **300** 确定的候选关键字字符串重复方框 **305**。在 **310**，至少部分地基于候选关键字字符串创建位向量。在 **315**，对应于候选关键字字符串的位向量被存储在数据库中。

现在转到图 4，从用户角度示出了根据本发明的一个实施例的用于关键字字符串的增量搜索的方法的流程图。在 **400**，包括一个或多个具有一个或多个符号的单词的输入关键字字符串被提交。如果存在任何匹配，在 **405** 至少一个候选关键字字符串被接收，其具有匹配于输入关键字字符串的位向量的位向量。根据本发明的一个实施例，在每次用户输入输入关键字字符串的符号时执行处理 **400**

和 **405**。用户可以继续输入用于搜索的符号，匹配的数量随着每个输入减少。

现在转到图 5，从搜索引擎的角度示出了根据本发明的一个实施例的用于关键字增量搜索的方法的流程图。在 **500**，接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串。在 **505**，至少部分地基于输入的关键字字符串创建位向量。在 **510**，位向量与表示候选关键字字符串的一个或多个位向量进行比较。根据本发明的一个实施例，位向量的比较与在关键字字符串中的关键字词首的顺序无关。例如，输入关键字字符串“JS”与候选关键字字符串“John Smith”和“Sam Jones”匹配。步骤 **510** 的结果是一组表示候选关键字字符串的匹配位向量。在 **515**，判断是否存在匹配。如果存在匹配，在 **520** 将常规关键字匹配算法应用于由匹配位向量组表示的候选关键字字符串。在 **525**，判断常规关键字匹配算法是否找到匹配。如果存在匹配，在 **535** 将匹配候选关键字呈现给用户。如果在 **515** 或 **525** 不存在匹配，则搜索在 **530** 停止。步骤 **505** 至 **515** 减少了常规关键字匹配算法必须运行的次数 (**520**)，因此提高了关键字字符串搜索的效率。

图 6A 至 6D 是示出了根据本发明的一个实施例的基于在输入关键字字符串中的一个或多个单词的第一符号创建位向量的方框图。图 6A 示出了基于输入字符串“Dances With Wolves” **600** 创建位向量。在“Dances” (**602**) 中的“D”对应于位 **608**。在“With” (**604**) 中的“W”和在“Wolves” (**606**) 中的“W”都对应于位 **610**。因此，表示输入字符串“Dances With Wolves” **600** 的位向量使位 **608** 和 **610** 被置位。

图 6B 示出了基于输入字符串“Jonathan Smith” **612** 创建位向量。“Jonathan” (**614**) 中的“J”对应于位 **618** 以及“Smith” (**616**)

中的“S”对应于位 **620**。因此，表示输入字符串“Jonathan Smith” **612** 的位向量使位 **618** 和 **620** 被置位。

图 6C 示出了基于输入字符串“Saint Louis” **622** 创建位向量。“Saint” (**624**) 中的“S”对应于位 **628** 和“Louis”中的“L”对应于位 **626**。因此，表示输入字符串“Saint Louis” **622** 的位向量使位 **626** 和 **628** 被置位。

图 6D 示出了基于输入字符串“Sierra Wireless”创建位向量。“Sierra” (**632**) 中的“S”对应于位 **636** 以及“Wireless” (**634**) 中的“W”对应于位 **638**。因此，表示输入字符串“Sierra Wireless” (**634**) 的位向量使位 **636** 和 **638** 被置位。

如在图 6A 至图 6D 中可以看到的，在该实施例中的位向量或掩码为用于英语语言的 26 位长，忽略了数字和大写字母/小写字母的区别。这些位向量可以在 32 位整数中表示，因此允许通过 32 位整数比较操作的关键字字符串比较，例如：

```

Int input_pattern          =
000100000000000000000001000B;

Int Dances_with_wolves   =
000100000000000000000001000B;

If ((input_pattern & dances_with_wolves) ==
input_pattern) {
//we have a potential match
}

```

如以上实例所示，为了可能的匹配评估整个关键字字符串要求相对较少数量的机器指令。通过将关键字模式置于相邻的存储器位置，因此避免指针加载或页交换，可以进一步优化机器指令时间。

现在转到图 7，示出了根据本发明的一个实施例的创建关键字字符串数据库的方法的流程图，该数据库包括具有用于字母表中每个符号的位位置的位向量，并且使对应于表示单词的第一符号的至少一个符号的位位置的位被置位的位向量。在 700，确定候选关键字字符串以存储在数据库中。每个候选关键字字符串执行方框 705。在 710，创建位向量，该位向量具有用于字母表中每个符号的位位置以及对应于表示候选关键字字符串中的单词的第一符号的至少一个符号的位位置的位被置位。在 715，对应于候选关键字字符串的位向量被存储。

根据本发明的一个实施例，字母表包括英语语言的字母表。根据本发明的实施例可以使用任何语言的字母表。

根据另一个实施例，还具有大写字母形式的符号共享位位置。根据本发明的另一个实施例，具有对应的重音形式的符号共享相同的位位置。例如，符号“Ä”、“Â”、“A”和“A”共享相同的位位置。

现在转到图 8，从使用图 7 的数据库的用户的角度示出了根据本发明的一个实施例的用于关键字字符串的增量搜索的方法的流程图。在 800，包括一个或多个包括一个或多个符号的单词的输入关键字字符串被提交。如果存在任何匹配，则在 805 匹配候选关键字字符串被接收，其中在每个候选关键字字符串中的至少一个单词的第一符号与输入关键字字符串中的对应单词的第一符号匹配。根据本发明的一个实施例，在每次用户输入输入关键字字符串的符号

时执行处理 **800** 和 **805**。用户可以继续输入用于搜索的符号，匹配的数量随着每次输入而减少。

现在转到图 9，从使用图 7 的数据库的搜索引擎的角度示出了根据本发明的一个实施例的用于关键字字符串的增量搜索的方法的流程图。在 **900**，接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串。在 **905**，创建具有用于字母表中每个符号的位位置的位向量、以及具有使对应于表示输入关键字字符串中的单词的第一符号的至少一个符号的位位置的位被置位的位向量。在 **910**，位向量与表示候选关键字字符串的一个或多个其它位向量进行比较。根据本发明的一个实施例，位向量的比较与在关键字字符串中的关键字词首的顺序无关。例如，输入关键字字符串“JS”与候选关键字字符串“John Smith”和“Sam Jones”匹配。步骤 **910** 的结果是一组表示候选关键字字符串的匹配位向量。在 **915**，判断是否存在匹配。如果存在匹配，在 **920** 将常规关键字匹配算法应用于由匹配位向量组表示的候选关键字字符串。在 **925**，判断常规关键字匹配算法是否找到匹配。如果存在匹配，则在 **935** 将匹配候选关键字呈现给用户。如果在 **915** 或 **925** 不存在匹配，则搜索在 **930** 停止。步骤 **905** 至 **915** 减少了常规关键字匹配算法必须运行的次数 (**920**)，因此提高了关键字字符串搜索的效率。

现在转到图 10，示出了以每个符号的相对使用频率为基础创建一个或多个符号的统计分组的方框图。表 **1000** 表示英语字母表中的符号的相对使用频率的判决结果。表 **1015** 示出了符号组合为符号组，其中每个组的使用频率等于组中每个成员使用频率的和。根据本发明的一个实施例，选择了一个分组使得组的使用频率之间的不同减到最小。

在图 10 中示出的频率表是用于示例的目的，并不是用于一任何方式进行限制。根据本发明的实施例，频率表可以基于一个或多

个说话方式，例如口头或书面语句。示例性的书面语句包括报纸、小说、论文、字典和例如讲演的口头说话方式的记录。本领域的普通技术人员将认识到可以使用其它的书面语句。一种或多种说话方式可以是任何语言、语言方言或俚语。同时可以至少部分地基于讲演者的人口统计状况（例如年龄或收入）选择一个或多个说话方式。同时可以至少部分地基于用户对无线用户装置的先前输入选择一个或多个说话方式。

现在转到图 11，示出了根据本发明的一个实施例的使用图 10 的统计分组基于输入关键字字符串中的关键字的第一符号创建位向量的方框图。如图 11 所示，如果位位置与包括输入字符串 **1100** 中的单词的第一符号匹配的符号的符号组相关联，位图 **1135** 中的一位被置位。“Dances With Wolves”**1100** 中的每个单词的第一符号是 “D” **1105**、“W” **1110** 和 “W” **1115**。符号 “D” **1105** 对应于位位置 **1125** 以及符号 “W” (**1110**, **1115**) 对应于位位置 **1120**，因此位 **1125** 和 **1120** 被置位。

现在转到图 12，示出了根据本发明的一个实施例的以符号统计分组中符号的成员资格为基础比较关键字字符串的方法的流程图。在 **1200**，一种语言中至少一个符号的相对使用频率被确定。根据本发明的一个实施例，相对使用频率包括语言中单词的第一字符中的符号的相对使用频率。在 **1205**，基于相对使用频率，向至少一个符号分配了统计权重。在 **1210**，至少一个符号中的每个被分配给多个组中的一个，使得用于包括每个组的符号的统计权重和之间的差别最小。在 **1215**，至少部分地根据第一关键字字符串的至少一个符号是否分配到与第二关键字字符串的至少一个对应符号相同的组，来比较第一关键字字符串和第二关键字字符串。

根据本发明的另一个实施例，以在固定分组或接近符号的相对使用频率的分组中的符号的成员资格为基础，来比较关键字字符

串。例如，可以根据在标准电话键盘上找到的符号分组（“ABC”、“DEF”、“GHI”、“JKL”、“MNO”、“PQRS”、“TUV”、和“WXYZ”）来对符号进行分组。那些本领域的普通技术人员应该明白其他分组是可能的。

现在转到图 13，示出了根据本发明的一个实施例的用于创建包括一个或多个候选关键字字符串的总位向量的关键字字符串数据库的方法的流程图。在 **1300**，确定候选关键字字符串以存储在数据库中。每个候选关键字字符串重复方框 **1305**。在 **1310**，创建位向量，其中每位对应于语言字母表中的一个或多个符号以及其中对应于候选关键字字符串中的单词的第一符号的位位置的位被置位。在 **1315**，对应于候选关键字字符串的位向量被存储。

现在翻转图 14，从使用图 13 的数据库的用户的角度示出了根据本发明的一个实施例的用于关键字字符串的增量搜索的方法的流程图。在 **1400**，提交包括一个或多个符号的输入关键字字符串，其中每个符号代表在搜索字符串中的单词的第一符号。如果存在任何匹配，则在 **1405** 一个或多个匹配候选关键字字符串被接收，其中每个候选关键字字符串中的第一符号被包括一个或多个符号的分组所包括。根据本发明的一个实施例，在每次用户输入输入关键字字符串的符号时执行过程 **1400** 和 **1405**。用户可以继续输入用于搜索的符号，匹配的数目随着每次输入而减少。

现在转到图 15，从使用图 13 的数据库的搜索引擎的角度示出了根据本发明的一个实施例的用于关键字字符串的增量搜索的方法的流程图。在 **1500**，接收包括一个或多个符号的输入关键字字符串，其中每个符号代表在搜索字符串的单词的第一符号。在 **1505**，创建位向量，其中每位对应于字母表中的一个或多个符号并且其中对应于候选关键字字符串中的单词的第一符号的位位置的位被置位。在 **1510**，位向量与表示候选关键字字符串的一个或多个其它位

向量相比较。根据本发明的一个实施例，位向量的比较与在关键字字符串中的关键字词首的顺序无关。例如，输入关键字字符串“JS”与候选关键字字符串“John Smith”和“Sam Jones”匹配。步骤 1510 的结果是一组表示候选关键字字符串的匹配位向量。在 1515，确定是否存在匹配。如果存在匹配，则在 1520 将常规关键字匹配算法应用于由匹配位向量组表示的候选关键字字符串。在 1525，确定常规关键字匹配算法是否找到匹配。如果存在匹配，则在 1535 将匹配候选关键字字符串呈现给用户。如果在 1515 或 1525 不存在匹配，搜索在 1530 停止。步骤 1505 至 1515 减少了常规关键字匹配算法必须运行的次数（1520），因此提高了关键字字符串搜索的效率。

现在转到图 16，示出了根据本发明的一个实施例的基于符号的统计分组中的关键字字符串的第一单词的第一符号的成员资格创建总位向量的方框图。如图 16 所示，“Dances”（1605）中的“D”对应于位向量 1650 的位 1625、“Dances”（1610）中的“a”对应于位向量 1655 的位 1630、“Dances”（1615）中的“n”对应于位向量 1660 的位 1635 以及“Dances”中的“c”对应于位向量 1665 的位 1640。将四个 8 位向量组合得到 32 位十六进制值 10400220（1645）。

现在转到图 17，示出了根据本发明的一个实施例的基于符号统计分组中的关键字字符串的第一单词的第一符号的成员资格创建总位向量的方框图。图 17 类似于图 16，除了相同的四个 8 位向量被用于在输入字符串中的每个单词的前四个符号。如图 17 所示，在“Dances”中的符号“D”（1702）、“a”（1704）、“n”（1706）和“c”（1708）分别对应于位 1728、1730、1738 和 1742。在“With”中的符号“W”（1710）、“i”（1712）、“t”（1714）和“h”（1716）分别对应于位 1726、1732、1736 和 1744。同样地，在“Wolves”中的符号“W”（1718）、“o”（1720）、“l”（1722）和“v”（1724）分别对应于位 1726、1734、1740 和 1744。四个 8 位向量的组合得到 32 位十六进制值 90580722（1746）。

现在转到图 18, 示出了根据本发明的一个实施例的用于创建关键字字符串数据库的方法的流程图, 该数据库包括一个或多个总位向量, 其中每位对应于字母表中的一个或多个符号, 并且其中对应于在候选关键字字符串的单词词首中的符号的位位置的位被置位。在 **1800**, 确定候选关键字字符串以存储在数据库中。每个候选关键字字符串重复方框 **1805**。在 **1810**, 创建位向量, 其中每位对应于字母表中的一个或多个符号以及其中对应于候选关键字字符串中单词词首中的符号的位位置的位被置位。在 **1815**, 对应于候选关键字字符串的位向量被存储。

现在转到图 19, 从使用图 18 的数据库的用户的角度示出了根据本发明的一个实施例的用于关键字字符串的增量搜索的方法的流程图。在 **1900**, 提交包括一个或多个包括一个或多个符号的单词的输入关键字字符串。在 **1905**, 一个或多个匹配候选关键字字符串被接收, 其中匹配候选关键字字符串的单词的词首包括至少一个符号, 该符号与输入关键字字符串中对应单词的对应符号属于相同的符号组。根据本发明的一个实施例, 在每次用户输入输入关键字字符串的符号时执行过程 **1900** 和 **1905**。用户可以继续输入用于搜索的符号, 匹配的数量随着每次输入而减少。

现在转到图 20, 从使用图 18 的数据库的搜索引擎的角度示出了根据本发明的一个实施例的用于关键字字符串的增量搜索的方法的流程图。在 **2000**, 接收包括一个或多个包括一个或多个符号的单词的输入关键字字符串。在 **2005**, 创建位向量, 其中每位对应于字母表中的一个或多个符号以及对应于候选关键字字符串中的单词的词首中的符号的位位置的位被置位。在 **2010**, 位向量与表示候选关键字字符串的一个或多个位向量相比较。根据本发明的一个实施例, 位向量的比较与在关键字字符串中的关键字词首的顺序无关。例如, 输入关键字字符串“J S”与候选关键字字符串“John Smith”和“Sam Jones”匹配。步骤 **2010** 的结果是一组表示候选关键字字符

串的匹配位向量。在 **2015**，确定是否存在匹配。如果存在匹配，则在 **2020** 将常规关键字匹配算法应用于由匹配位向量组表示的候选关键字字符串。在 **2025**，确定常规关键字匹配算法是否找到匹配。如果存在匹配，则在 **2035** 将匹配候选关键字呈现给用户。如果在 **2015** 或 **2025** 不存在匹配，则搜索在 **2030** 停止。步骤 **2005** 至 **2015** 减少了常规关键字匹配算法必须运行的次数 (**2020**)，因此提高了关键字字符串搜索的效率。

现在转到图 21，示出了根据本发明的一个实施例的将总位向量的至少一位分配给至少一个关联于关键字字符串的非字符属性的方框图。在数据库中的联系人具有例如有线电话号码 (**2110**)、移动电话号码 (**2100**)、电子邮件地址 (**2105**) 和纸件邮件地址 (**2115**) 的属性。本领域的普通技术人员将意识到许多其它属性可以以类似的方式存储。

服务可能具有许多可以以类似的方式表示的非字母属性。不同旅馆和餐厅的价格可以以一至五位美元定级 (\$至\$\$\$\$\$)。另外，旅馆和餐厅的品质可以以一至五个星定级 (*至*****)。作为另一个实例，餐厅可以根据烹调风格类型分类 (例如，法国式、意大利式、美国式)。根据本发明的一个实施例，非字母属性被过载。根据本发明的另一个实施例，非字母属性以统计地平衡其使用量的方式被过载，例如如上所述的关于字母的属性。根据本发明的另一个实施例，相互之间具有强关联的属性被过载。例如，如果在服务价格定级和服务质量定级之间具有强关联，用来表示具有强关联的属性的位被过载。参考下面的图 22 示出。

现在转到图 22，示出了根据本发明的一个实施例的将总位向量的至少一位分配给至少一个关联于关键字字符串的非字符属性的方框图。如图 22 所示，位 **2234** 至 **2242** 用于价格定级和质量定级两者。例如，通过将位 **2234** 的值置位为 “1” 来表示 “\$” 的价格

定级。通过将相同的位 (**2234**) 置位为相同的值 (1) 来表示 “*” 的质量定级。

根据本发明的另一个实施例, 如果在预定的时间之后搜索没有完成, 则在此描述的搜索算法被取代。根据本发明的一个实施例, 预定的时间为两秒钟。

根据本发明的另一个实施例, 关键字字符串按层次组织。首先输入关键字字符串与顶层相关联的关键字字符串进行匹配。如果存在匹配, 则匹配的关键字从输入关键字字符串中除去, 然后接着关联于与匹配关键字相关联的部分层次的关键字字符串进行匹配。该过程将持续到没有匹配为止或到达层次的底部为止。下面参考图 23A 至 25 进行更详细的解释。

图 23A 至 23C 是示出了根据本发明的一个实施例的关键字字符串的分层组织的方框图。图 23A 至 23C 的层次代表用于北美地区的关键字字符串。参考图 23A, 最高级或顶级节点 **2300** 代表北美。节点 **2300** 的下一级或子节点代表在北美中的两个地区: 加拿大 **2302** 和美国 **2308**。在加拿大 **2302** 中的地区由子节点 **2304** 和 **2306** 表示。在美国 **2308** 中的地区由子节点 **2310** 和 **2312** 表示。根据本发明的一个实施例, 如果输入关键字字符串的每个单词与在导向特定分层元素的分层中的单词的词首匹配, 则输入关键字字符串与特定分层元素关键字字符串匹配。在图 23A 至 23C 示出了三个实例。如图 23A 所示, 用户可以通过输入输入关键字字符串 “c a” **2342** 来将地区指定在加拿大。关键字字符串 **2342** 中的 “a” 与 “North America” **2300** 中的词首 “a” 匹配。关键字字符串 **2342** 中的 “c” 与 “Canada” 中的词首 “c” 匹配。节点 **2304** 和 **2306** 是节点 **2302** 的子节点, 所以在图 23A 的分层中与输入字符串匹配的关键字字符串是 “British Columbia” **2304** 和 “Alberta” **2306**。

参考图 23B, 输入关键字字符串 “co a” 2344 与 “British Columbia” 相匹配, 这是因为输入关键字字符串 2344 中的 “a” 与 “North America” 2300 中的词首 “A” 匹配并且输入关键字字符串 2344 中的 “co” 与 “British Columbia” 中的词首 “Co” 匹配。

参考图 23C, 输入关键字字符串 “u y” 2346 与 “New York” 相匹配, 这是因为输入关键字字符串中的 “u” 与 “United States” 中的词首 “U” 匹配并且输入关键字字符串 2346 中的 “y” 与 “New York” 中的词首 “y” 匹配。

现在转到图 24, 示出了根据本发明的一个实施例的用于层次结构的关键字字符串的增量搜索的方法的流程图。在 2400, 接收包括一个或多个包括一个或多个符号的关键字的输入关键字字符串或模式。在 2405, 接收层次。层次的元素包括代表关键字字符串的中间节点和叶节点。在 2410, 计算由在层次中的每个中间节点和叶节点代表的关键字字符串的位向量。在 2415, 匹配项组被初始化以指示没有匹配项。在 2420, 为了匹配搜索在过程 2410 中创建的位向量层次。在 2425, 将匹配项输出或呈现给用户。

现在转到图 25, 示出了根据本发明的一个实施例的用于为了与模式匹配而搜索位向量的层次的方法的流程图。图 25 提供了图 24 中附图标号 2420 的更多细节。在图 24 中创建的位向量层次中的每个元素重复图 25 所示的过程。在 2500, 输入关键字字符串或模式被保存。在 2505, 确定对模式和元素执行的逻辑 “和” 操作的结果是否为 0。如果结果为 “否”, 则那些作为元素中单词词首的模式单词从模式中除去。在 2515, 确定模式是否为空。如果模式为空, 则在 2520 元素被添加到匹配项列表。在 2525, 确定模式是否为层次。如果模式为层次, 则对模式和元素执行图 25 中示出的搜索过程。在 2535, 保存的输入关键字字符串或模式被恢复。

虽然示出和描述了本发明的实施例和应用，显然，在没有背离本发明思想的前提下，由于本公开而受益的本领域的技术人员可以做出许多不同于上述的修改。因此，本发明除了在权利要求的精神之外，不被限制。

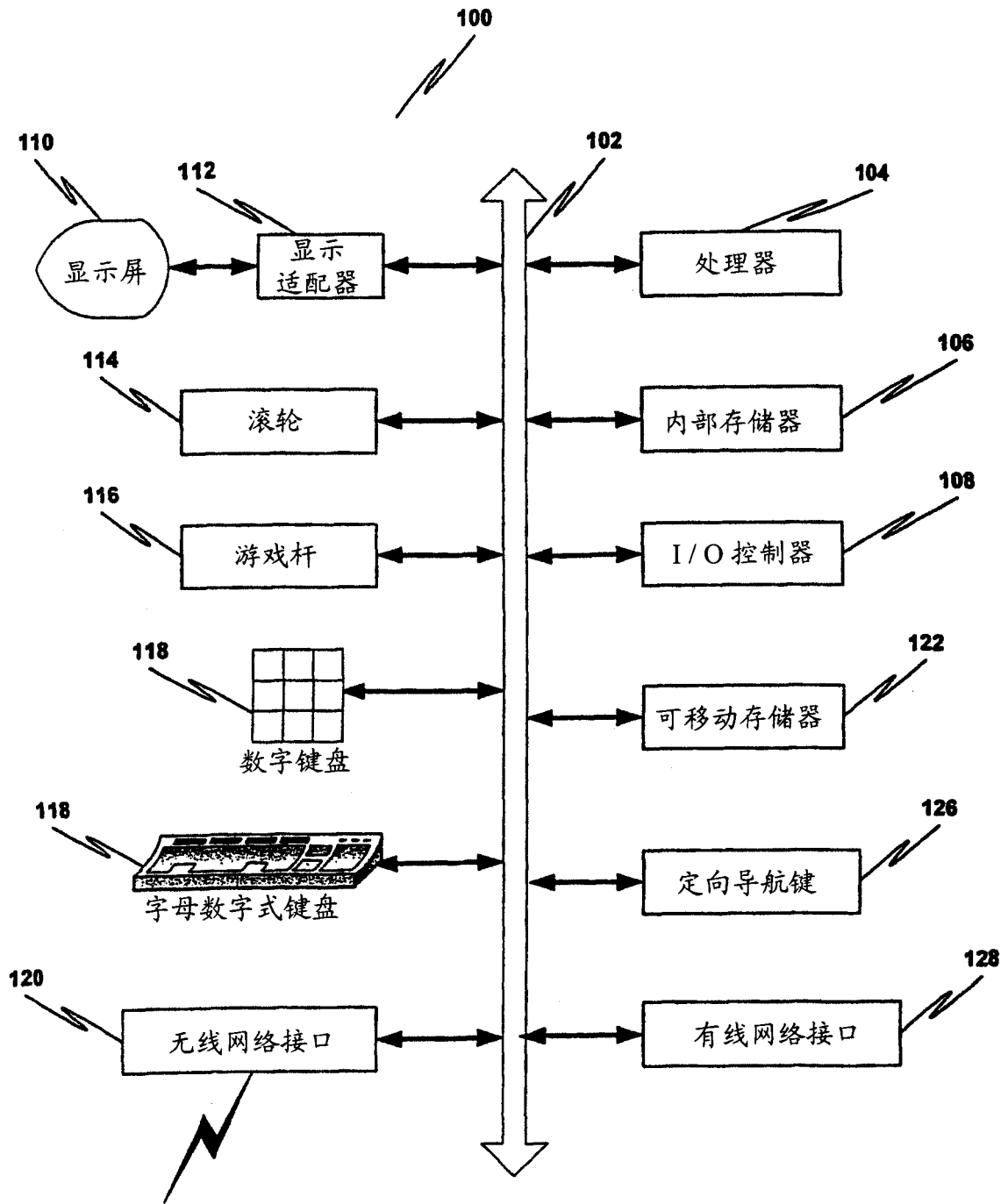


图 1

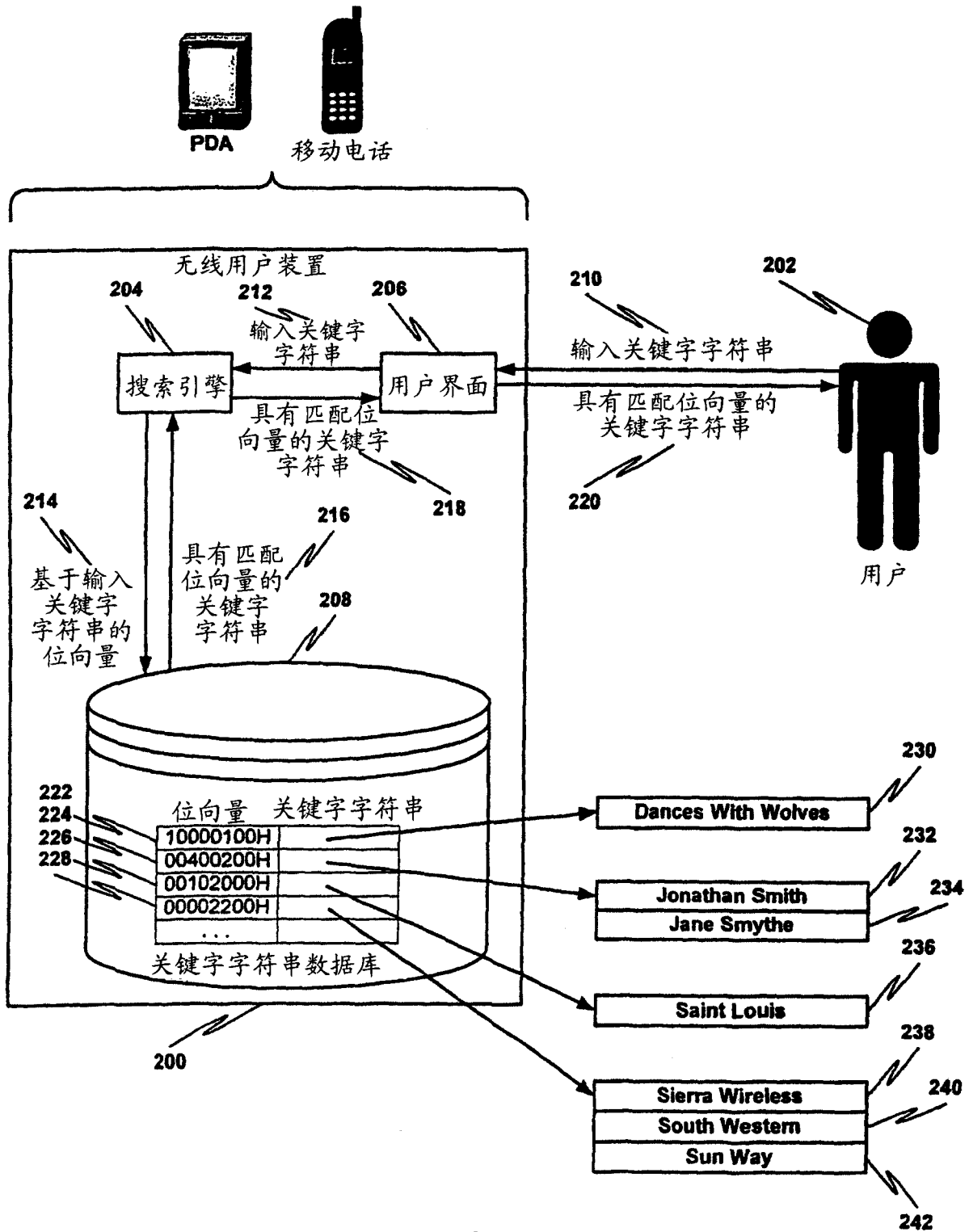


图2

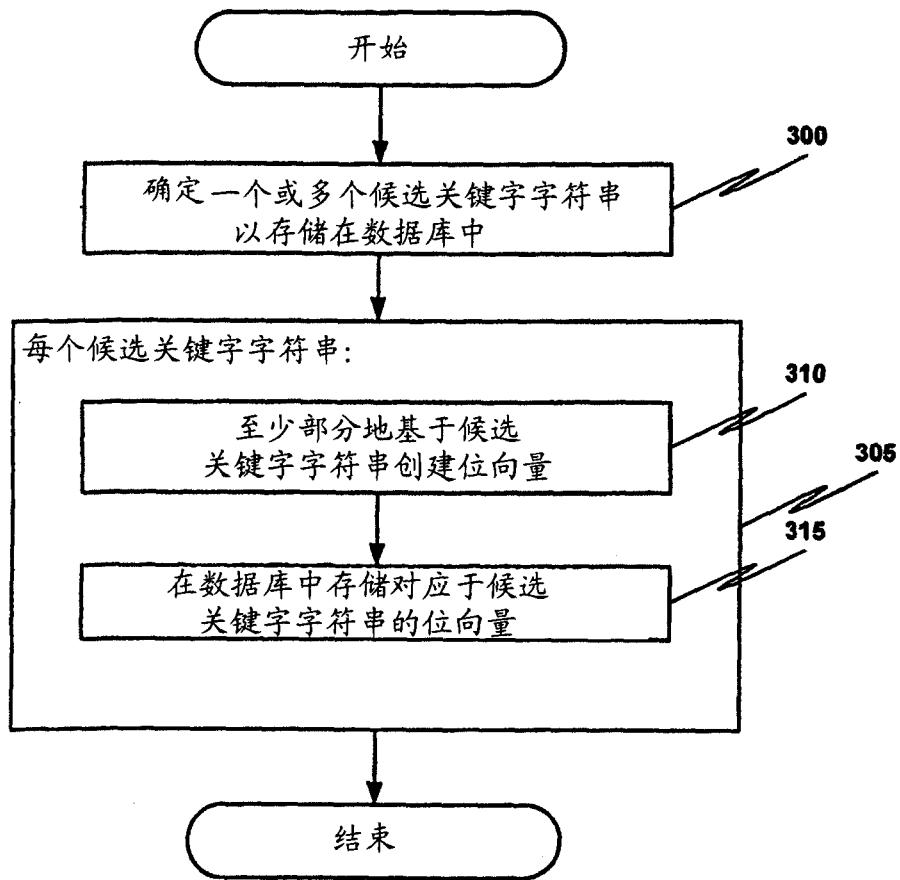


图3

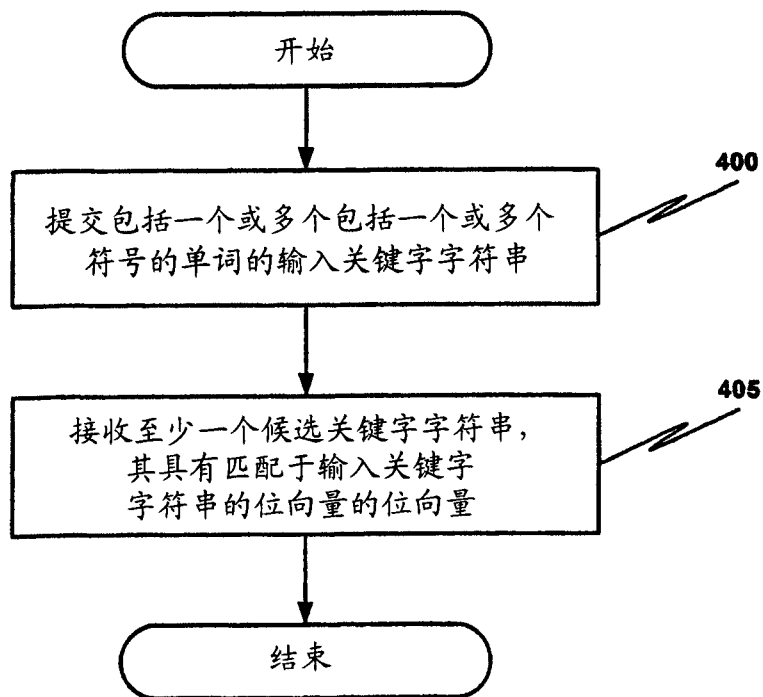


图4

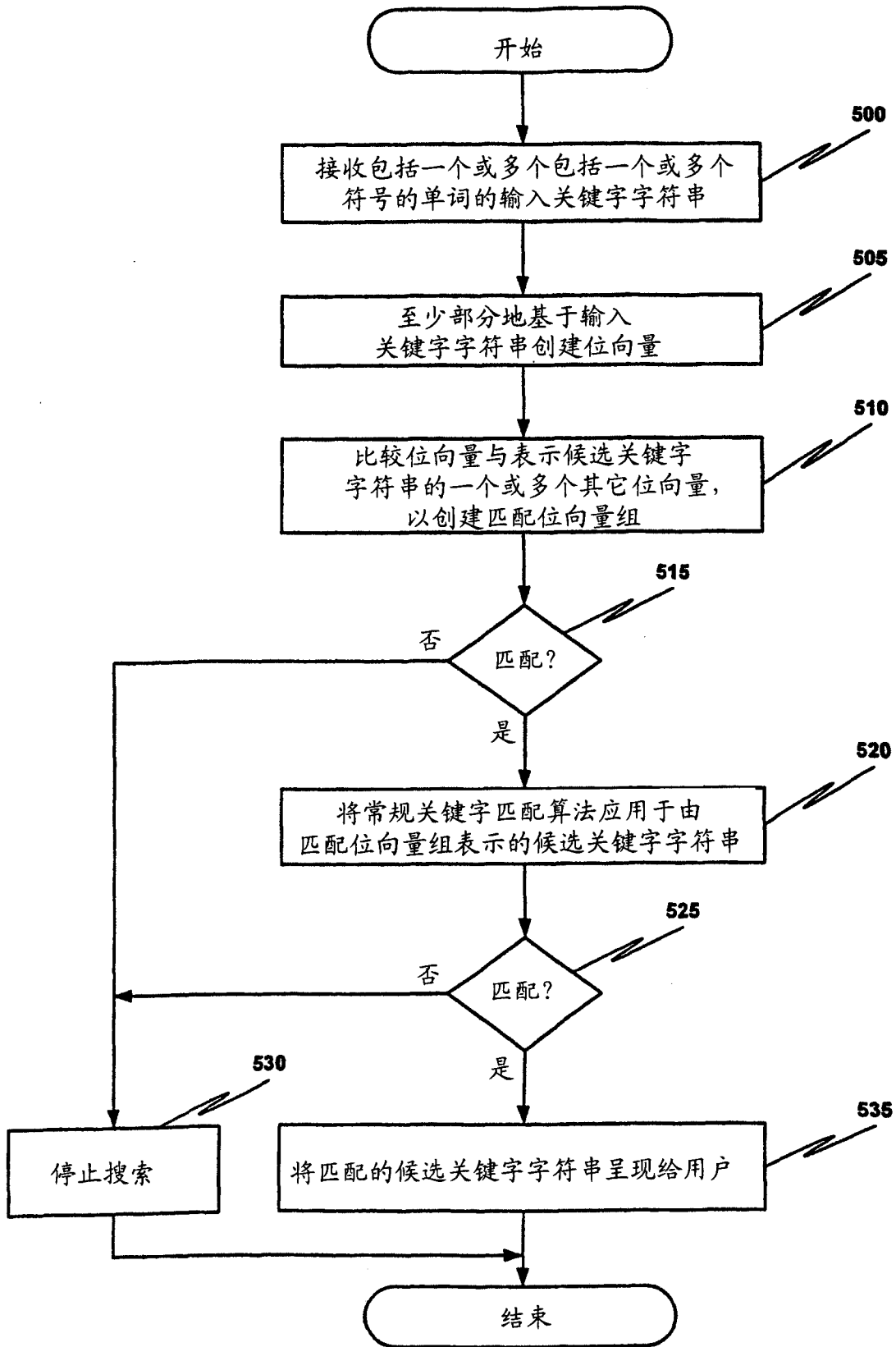
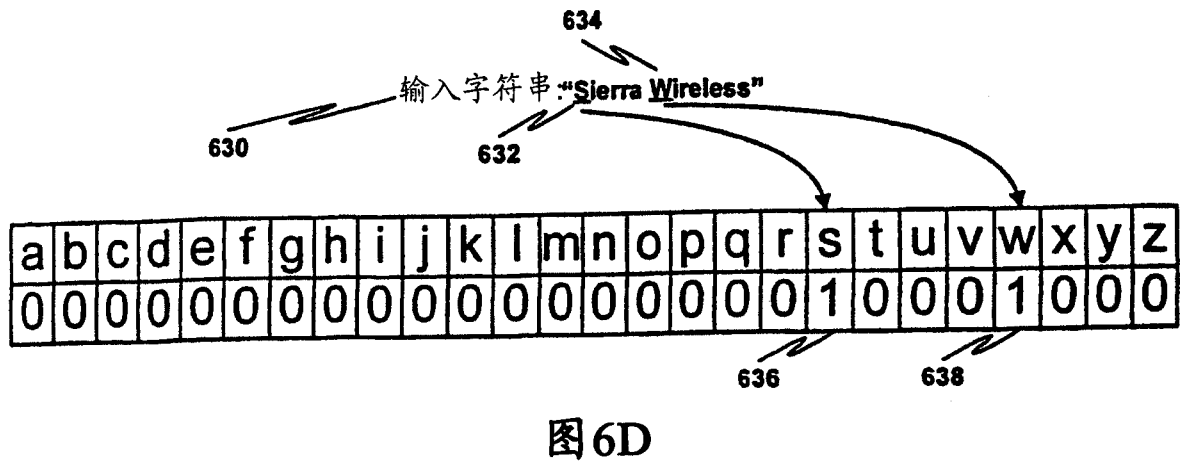
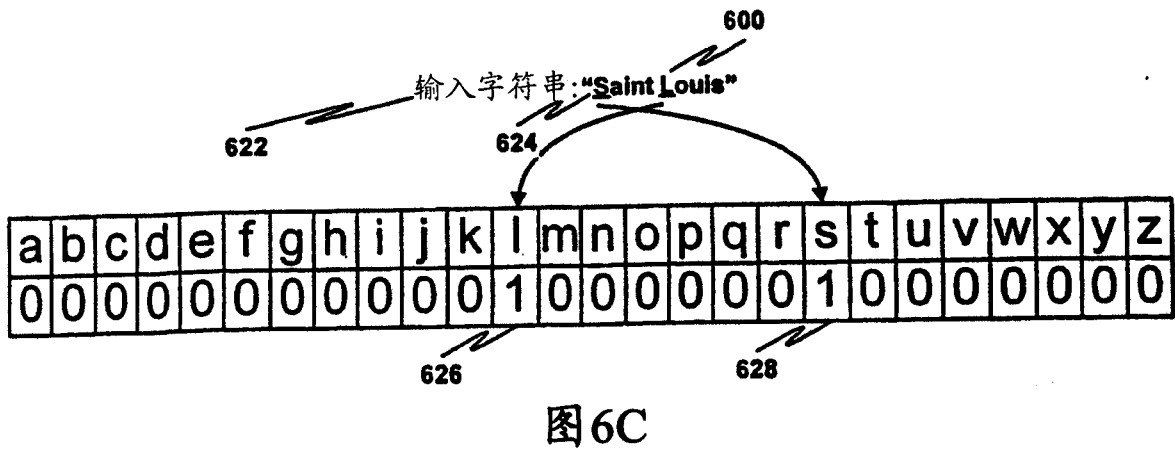
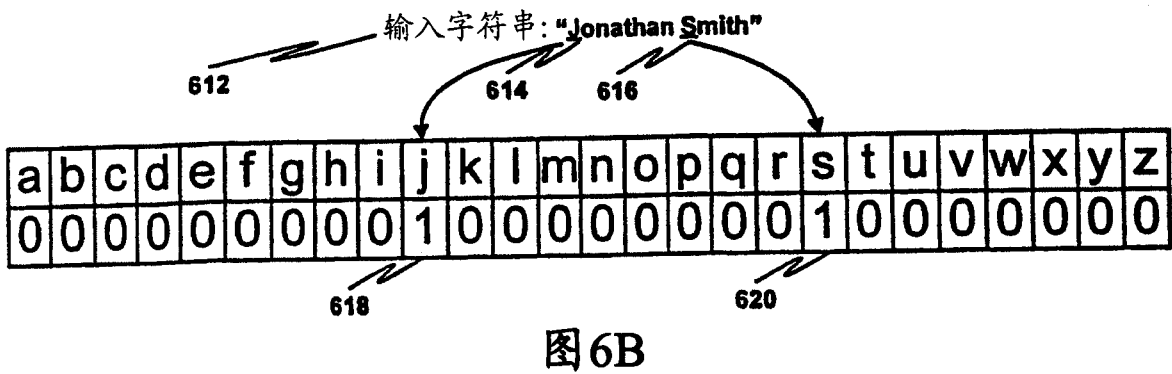


图5



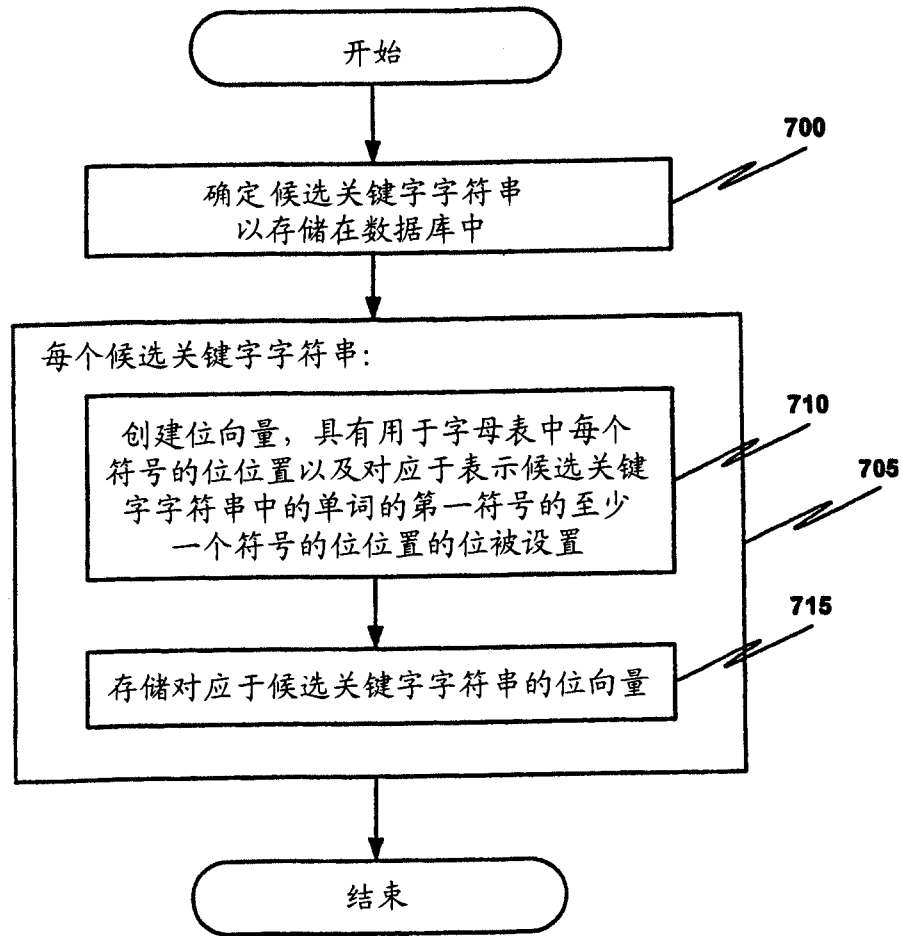


图7

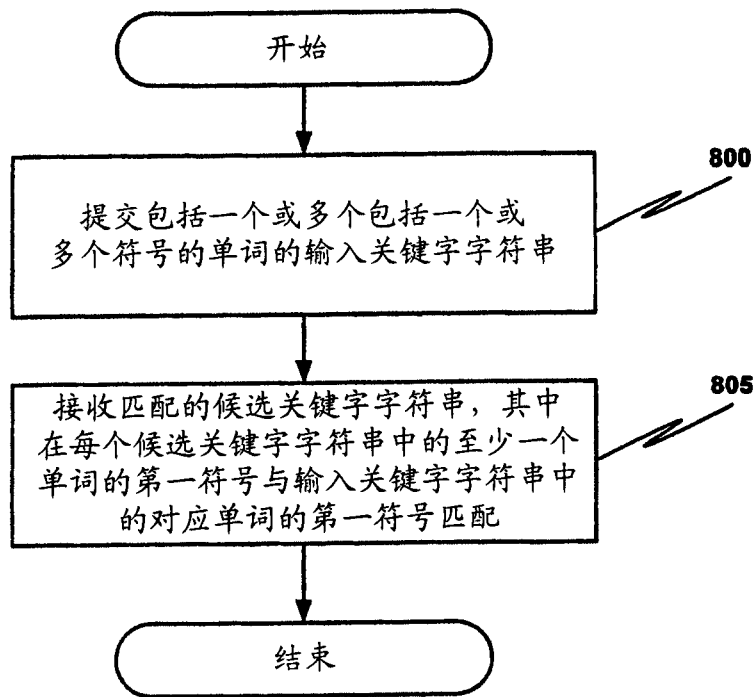


图8

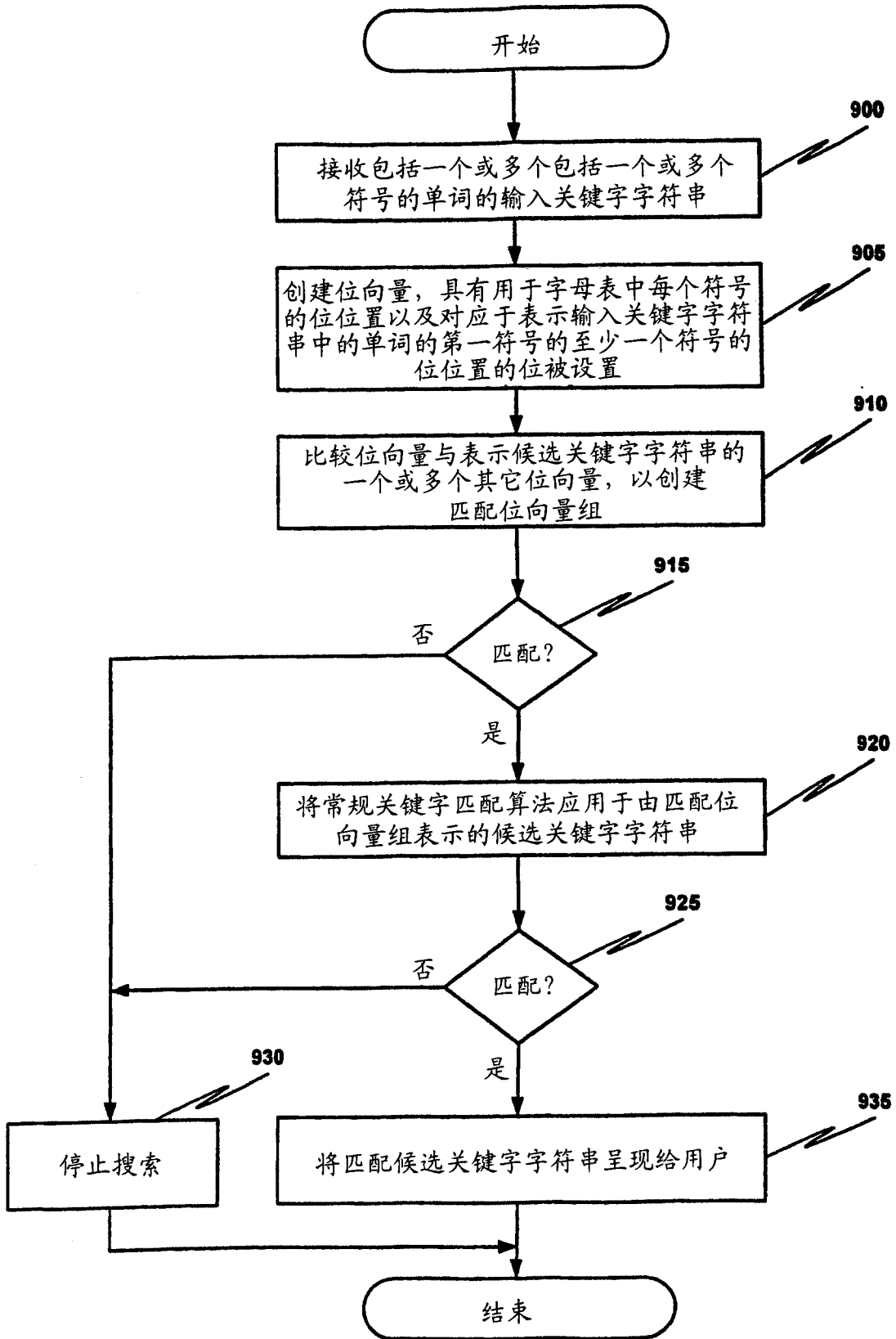


图9

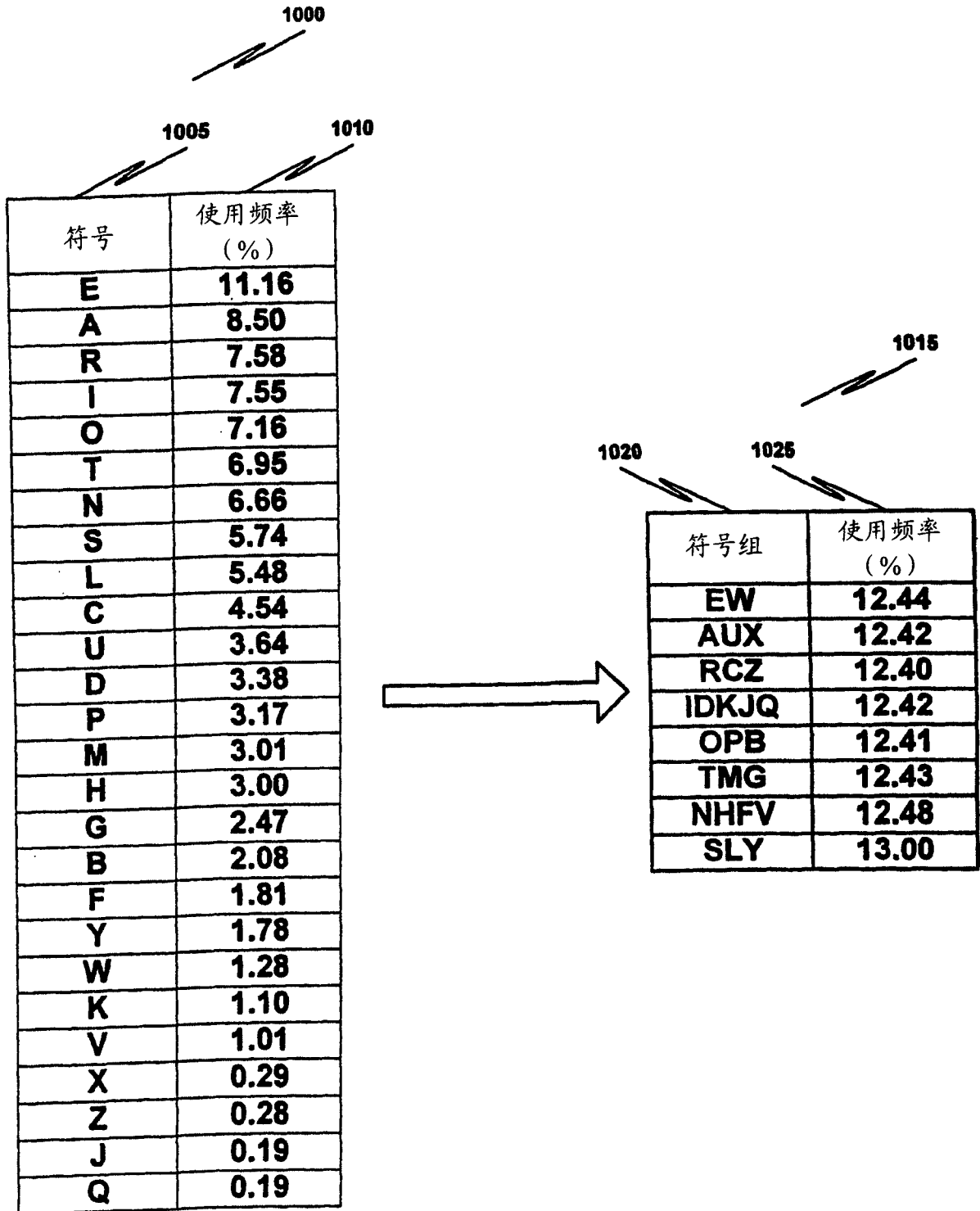


图10

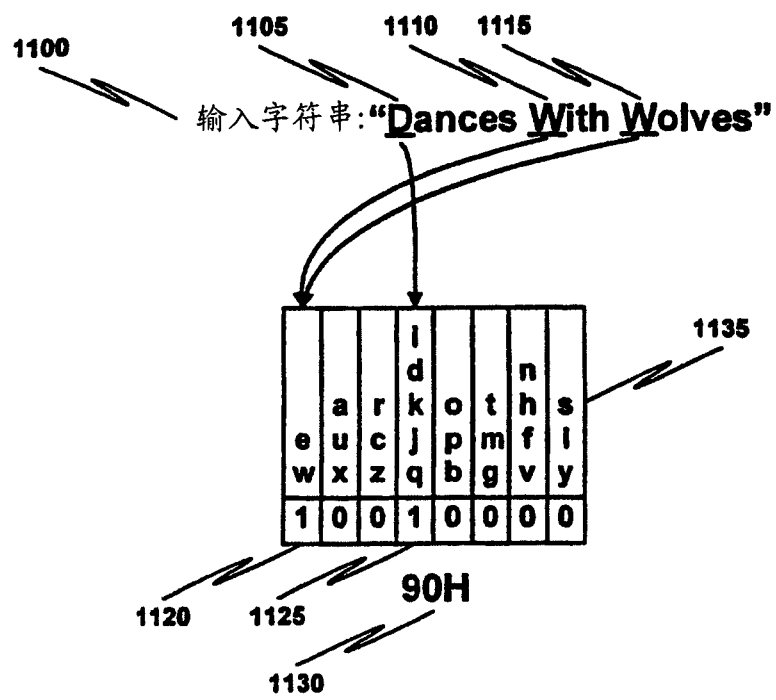


图 11

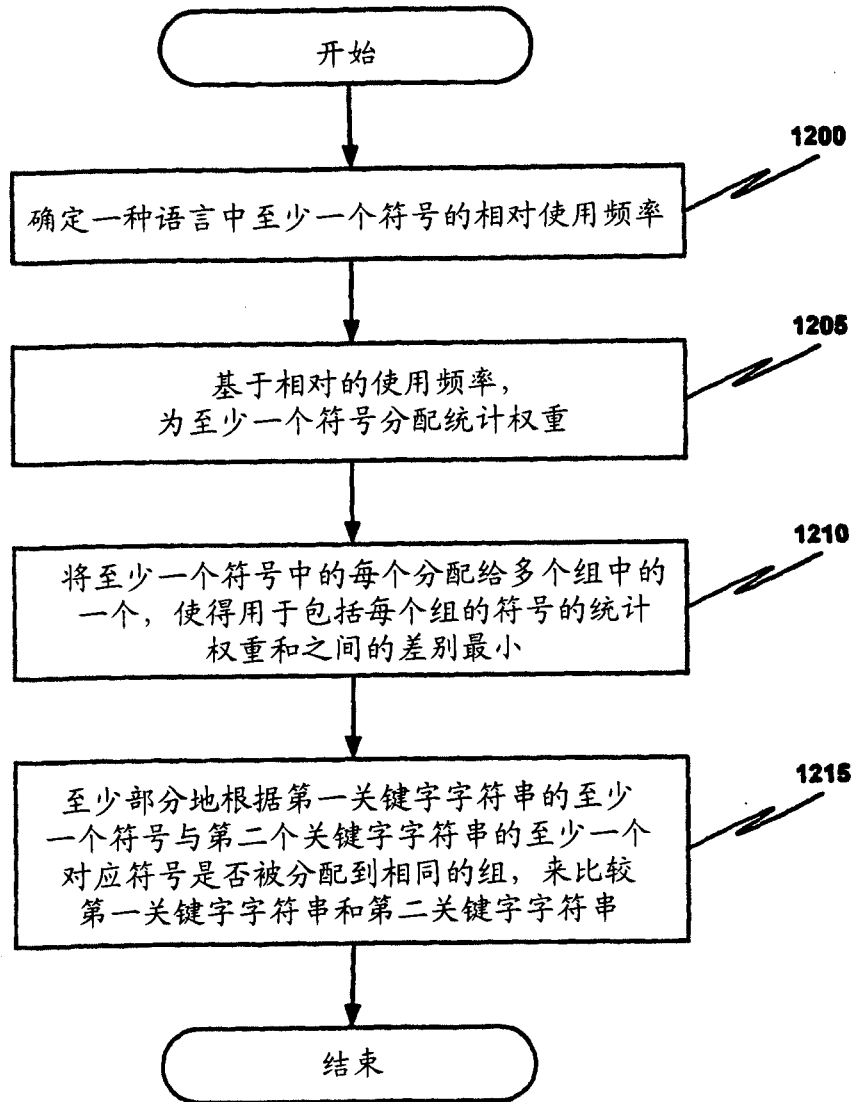


图12

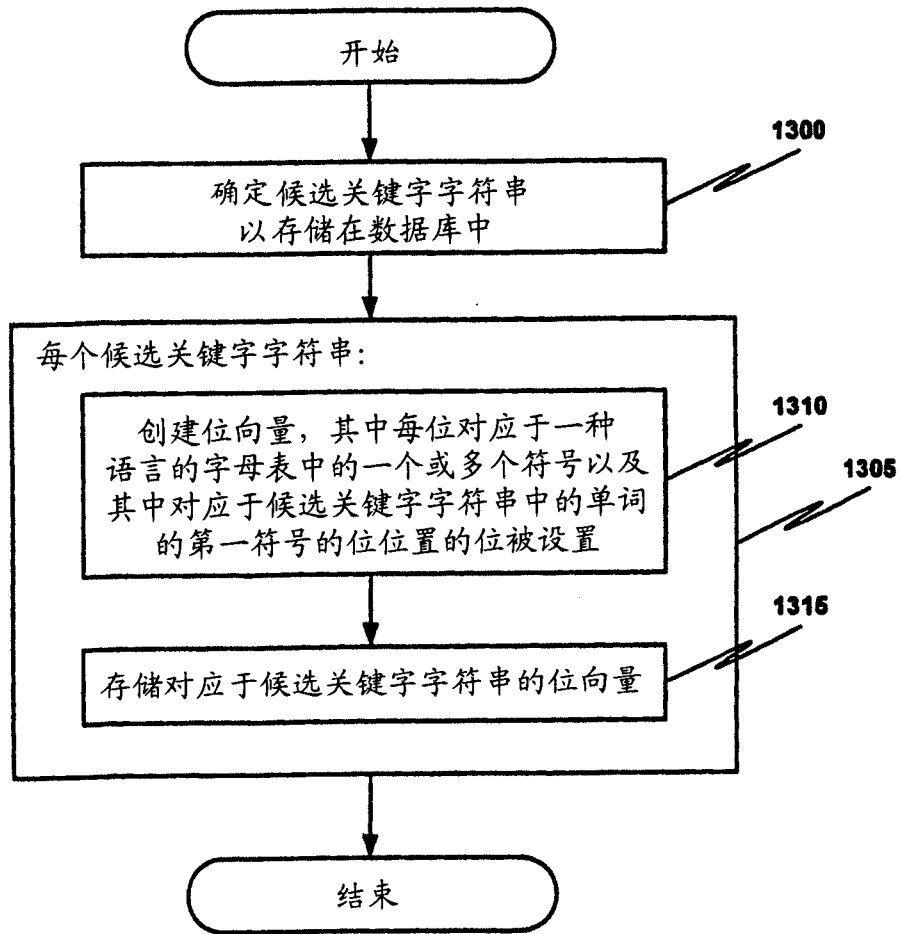


图13

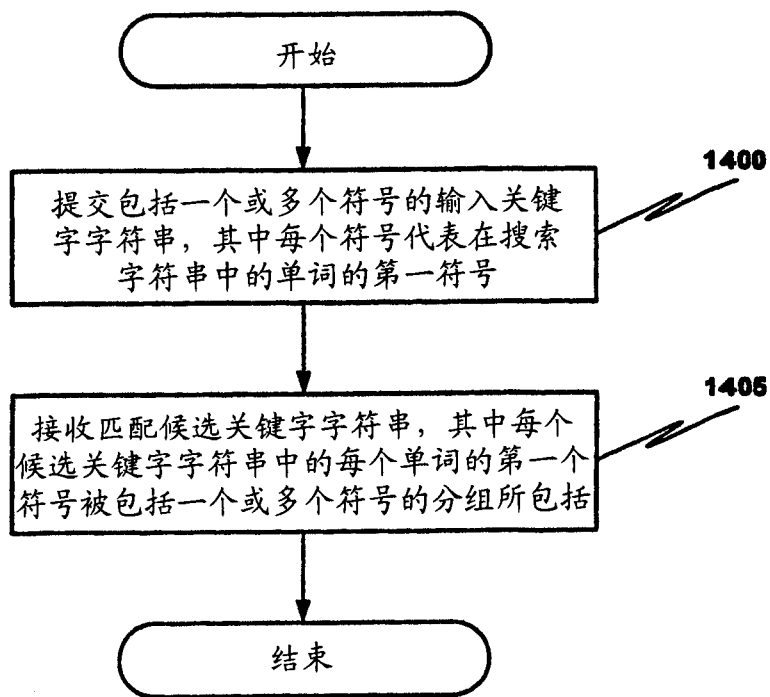


图 14

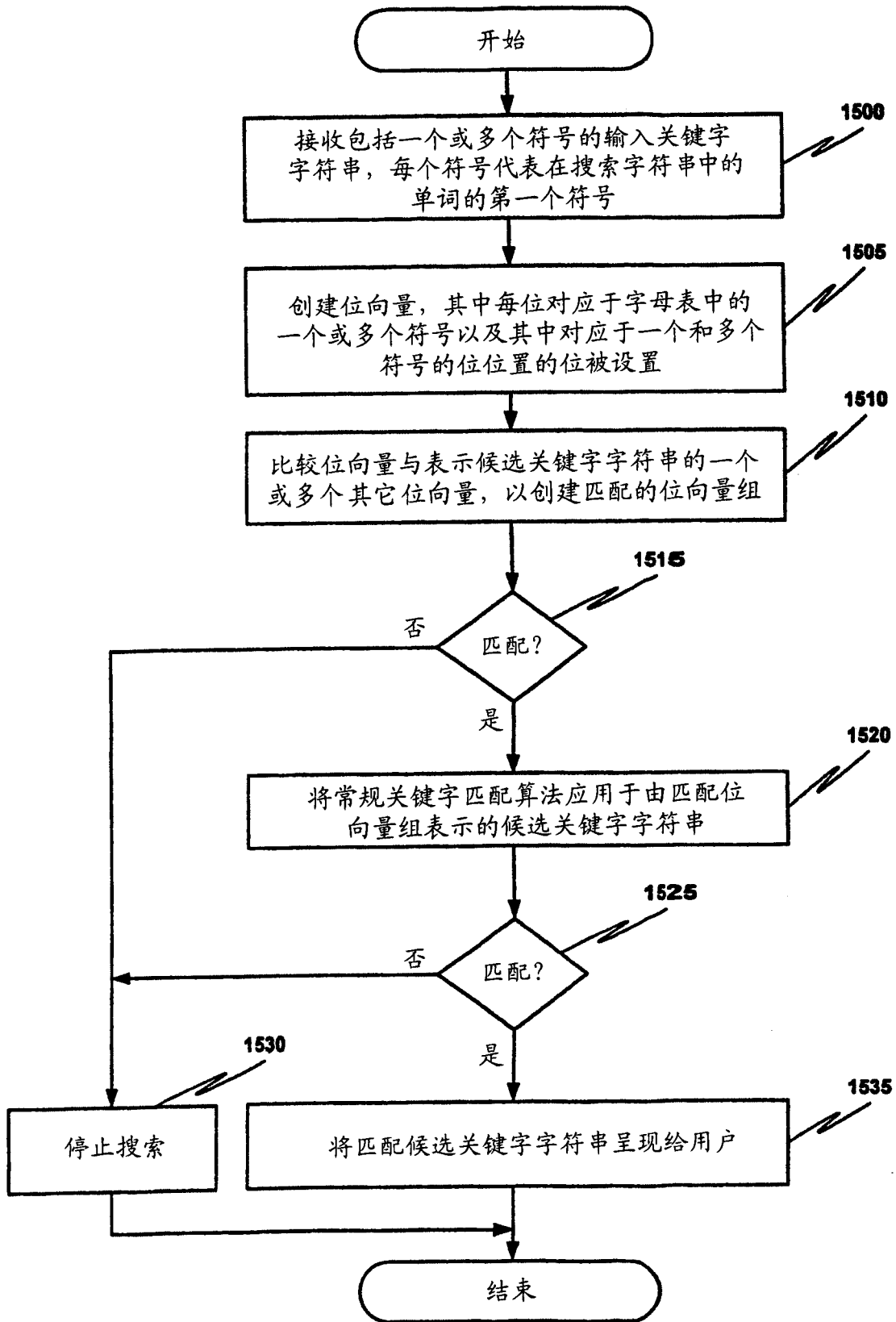


图15

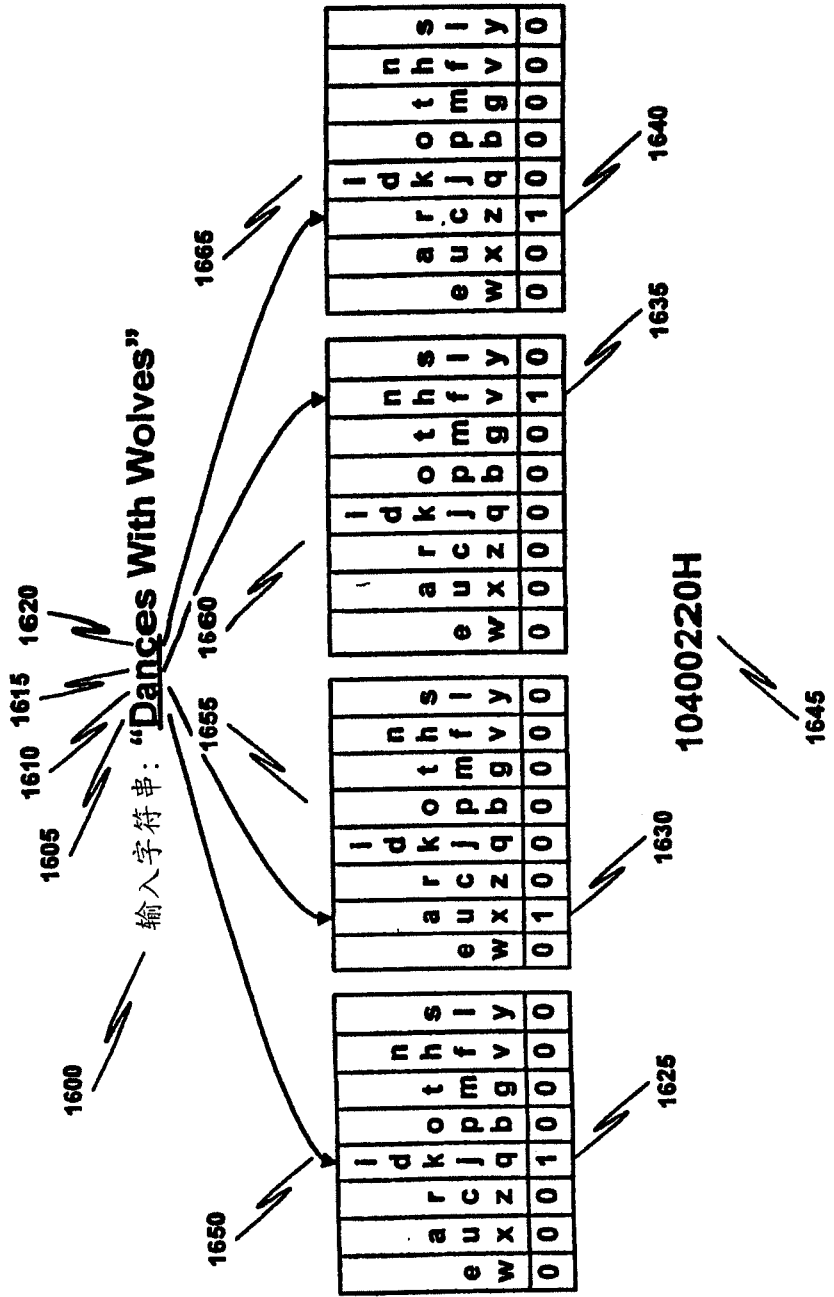


图16

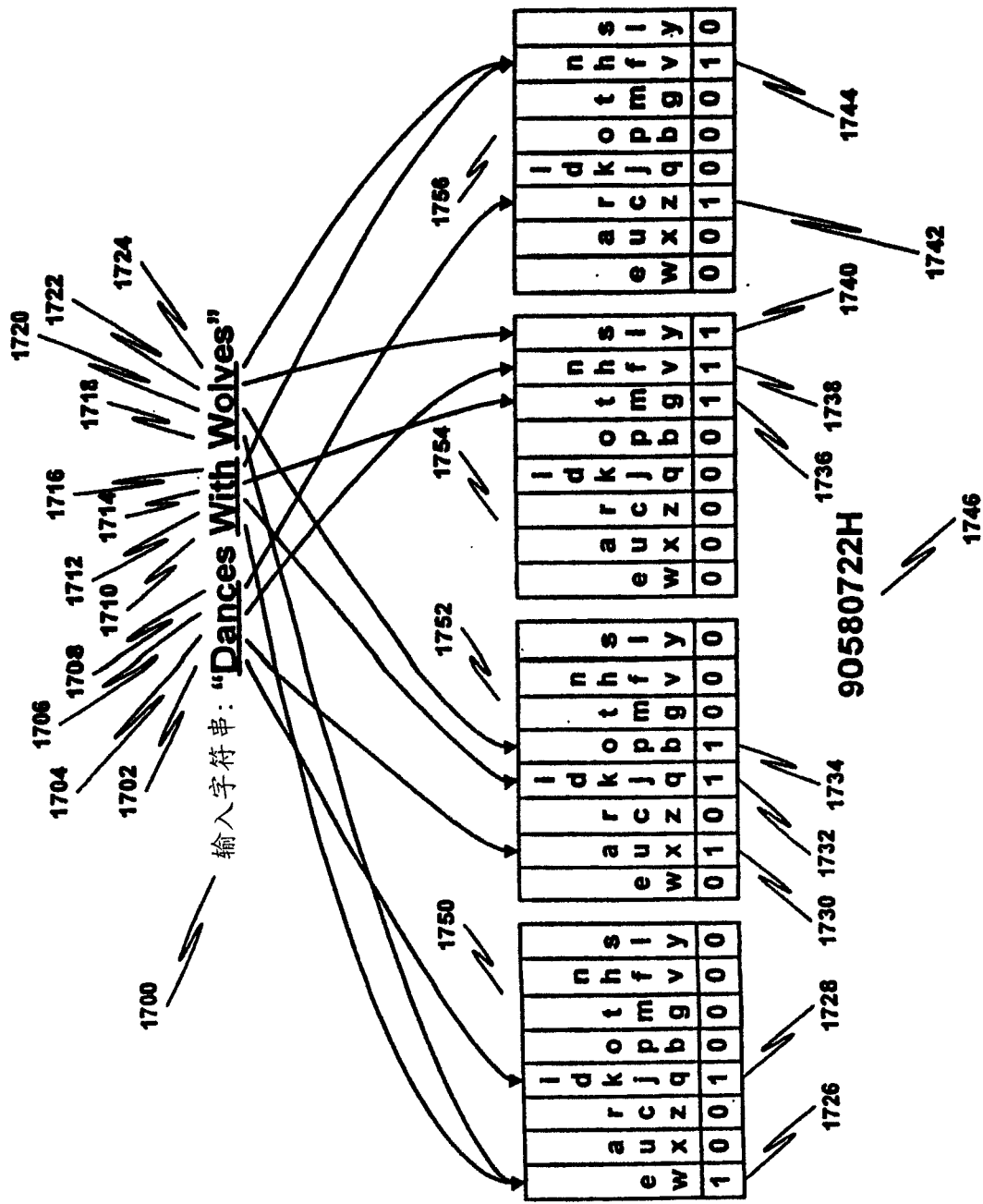


图17

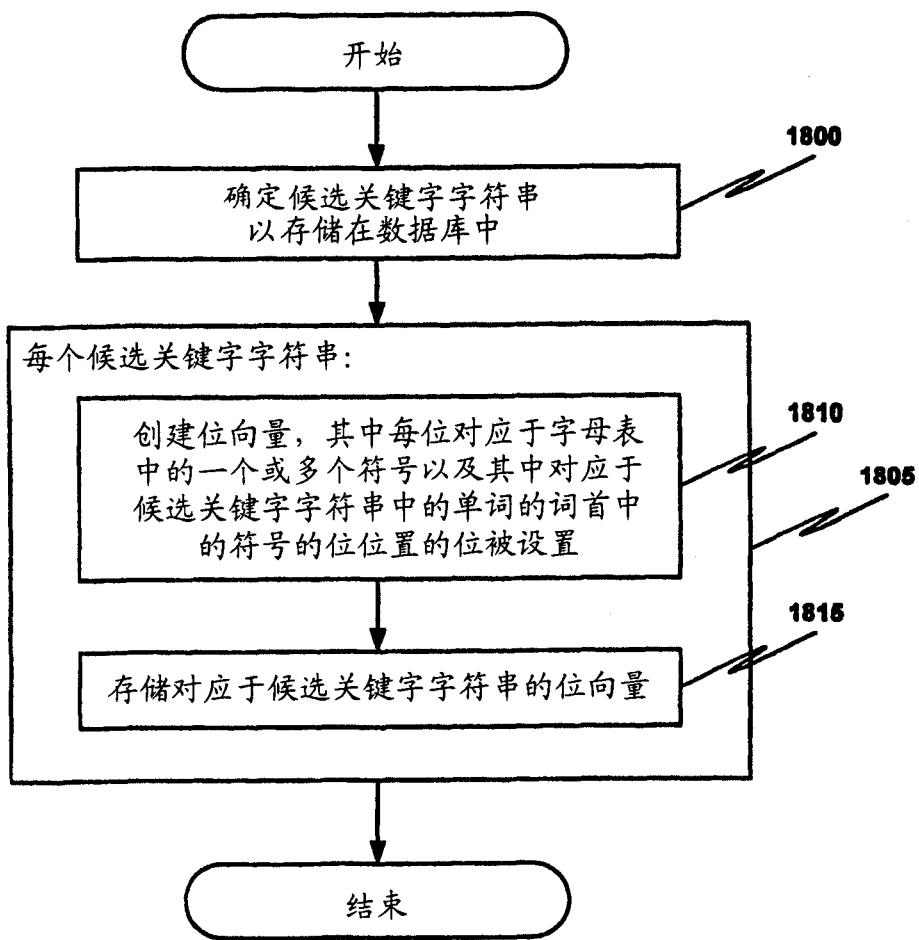


图18

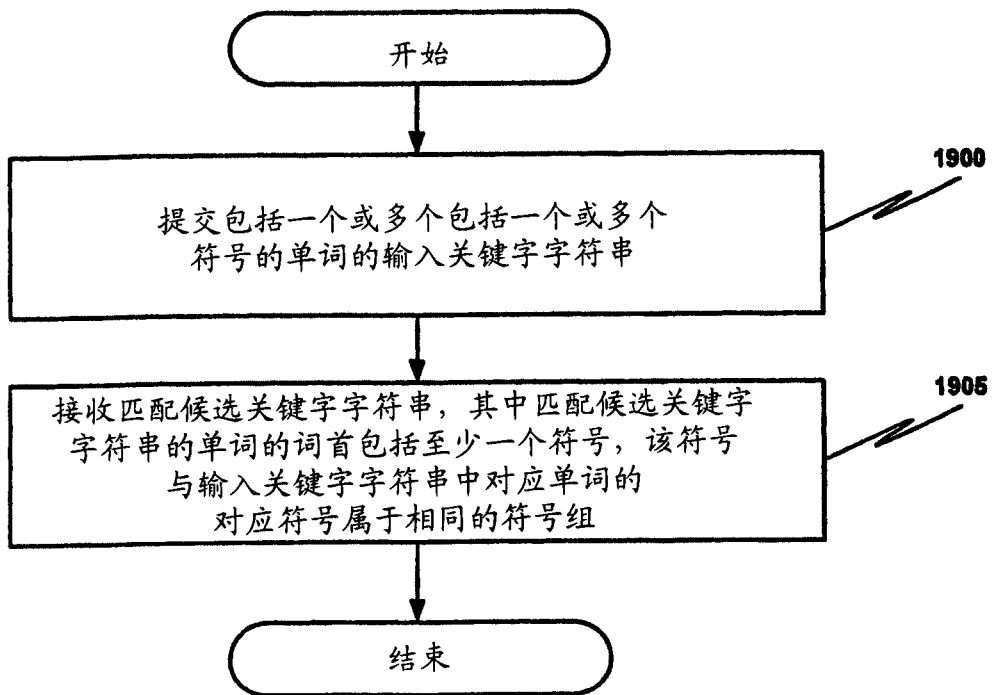


图19

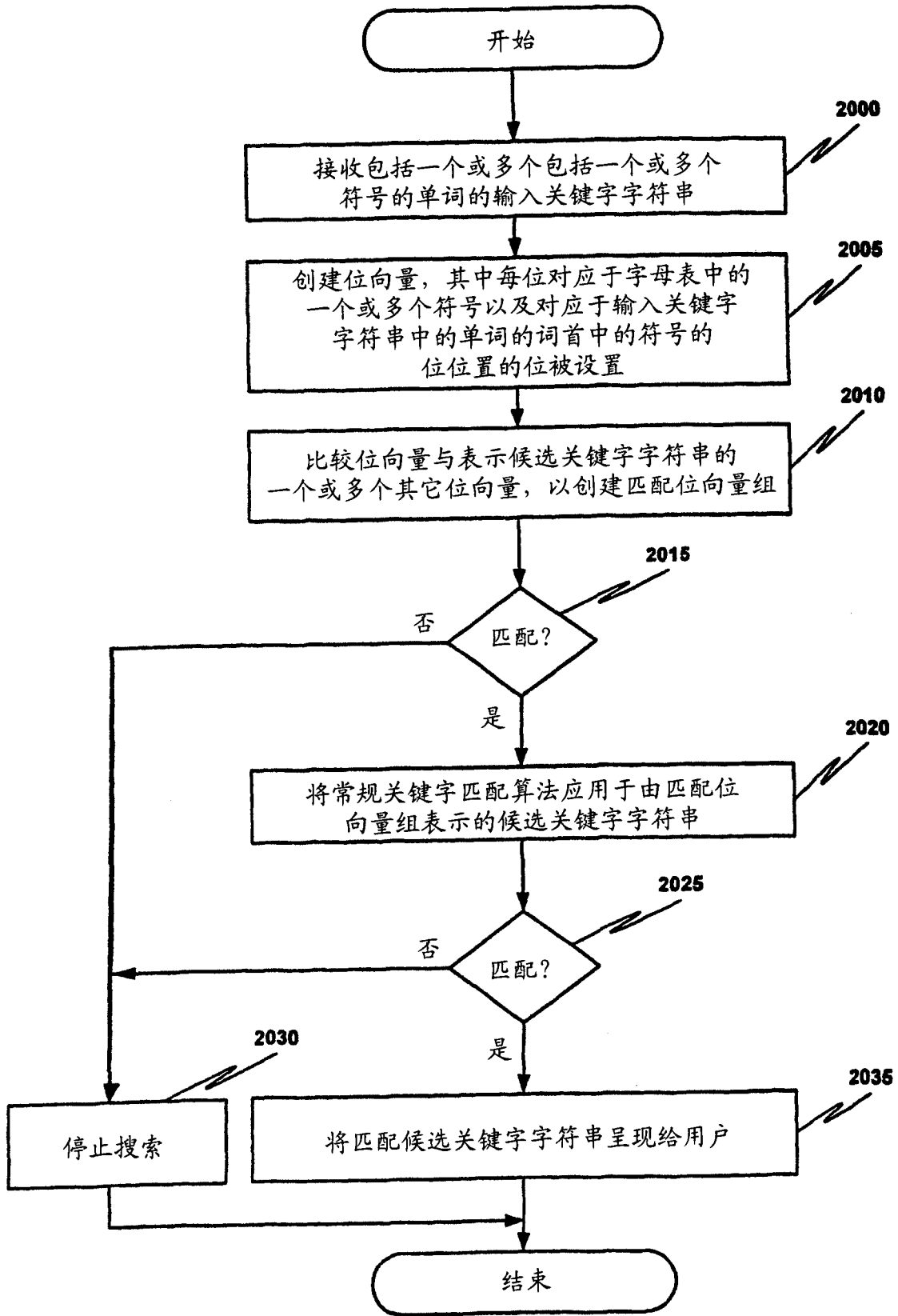


图20

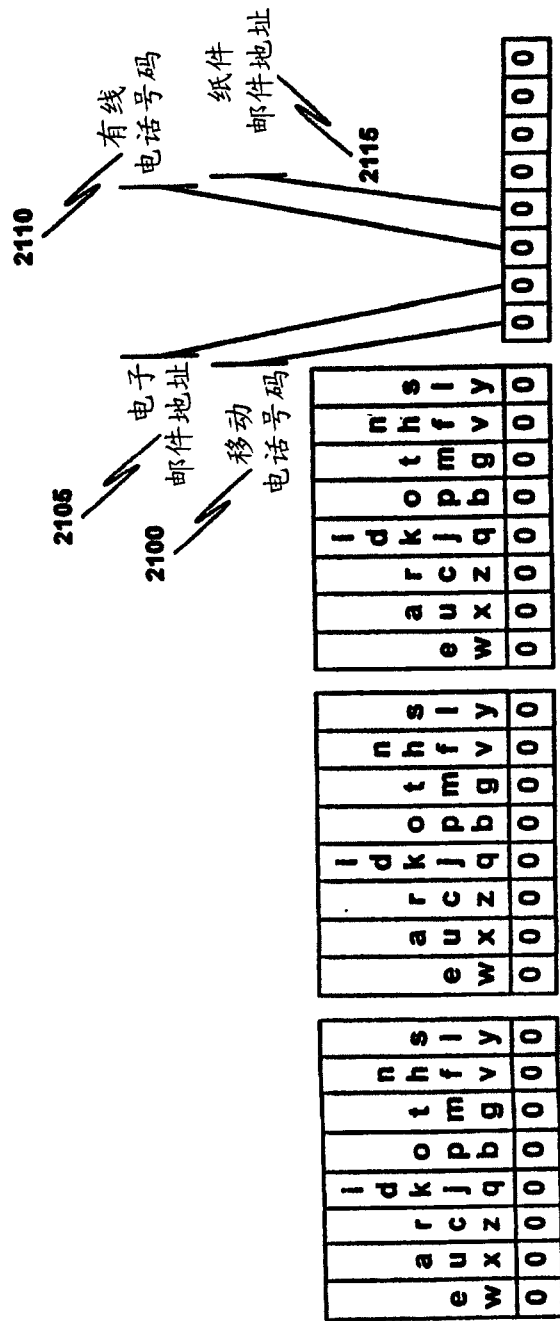


图21

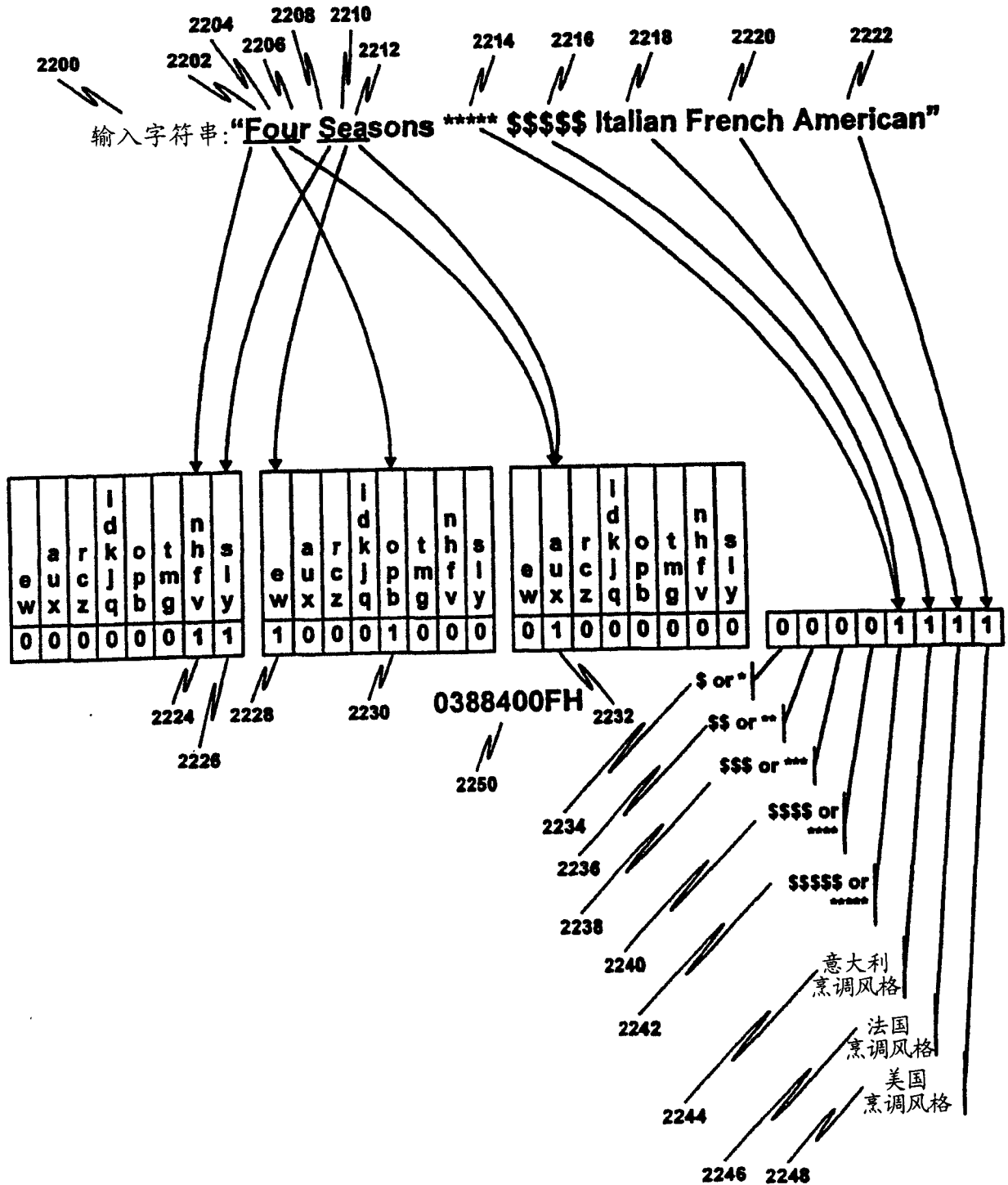


图22

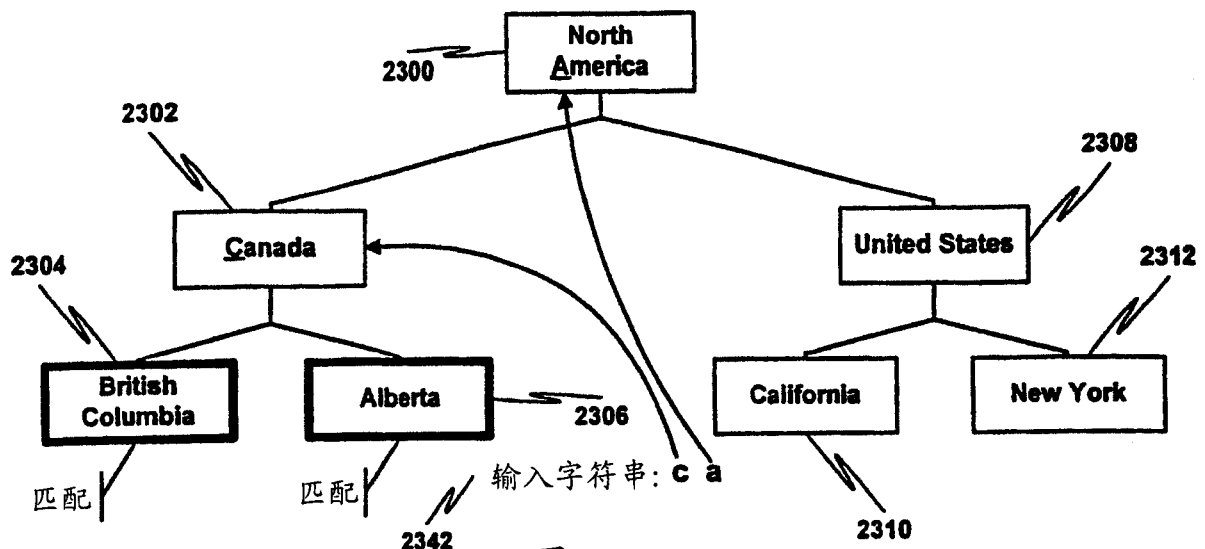


图23A

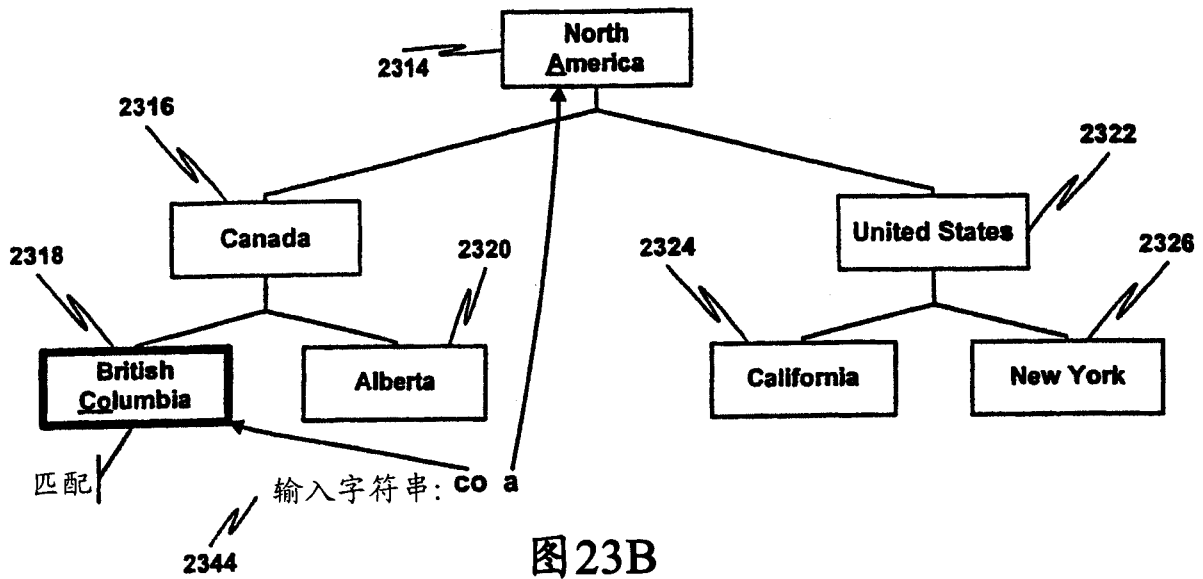


图23B

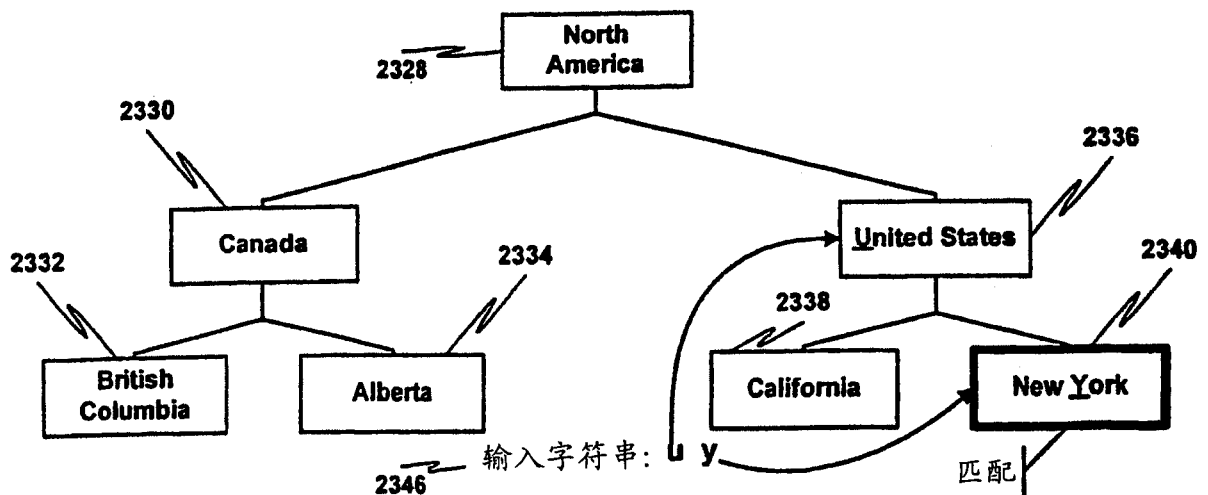


图23C

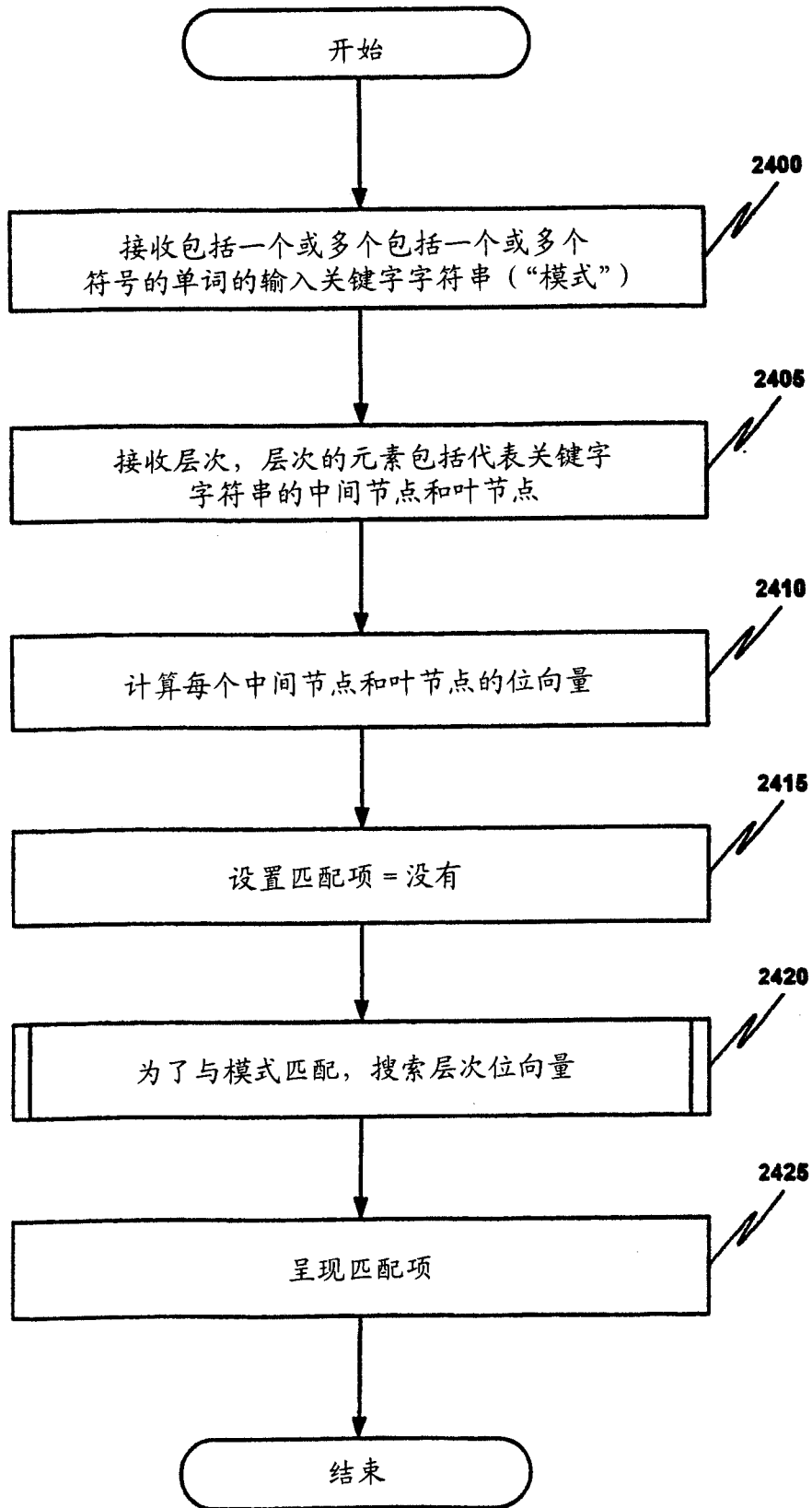


图24

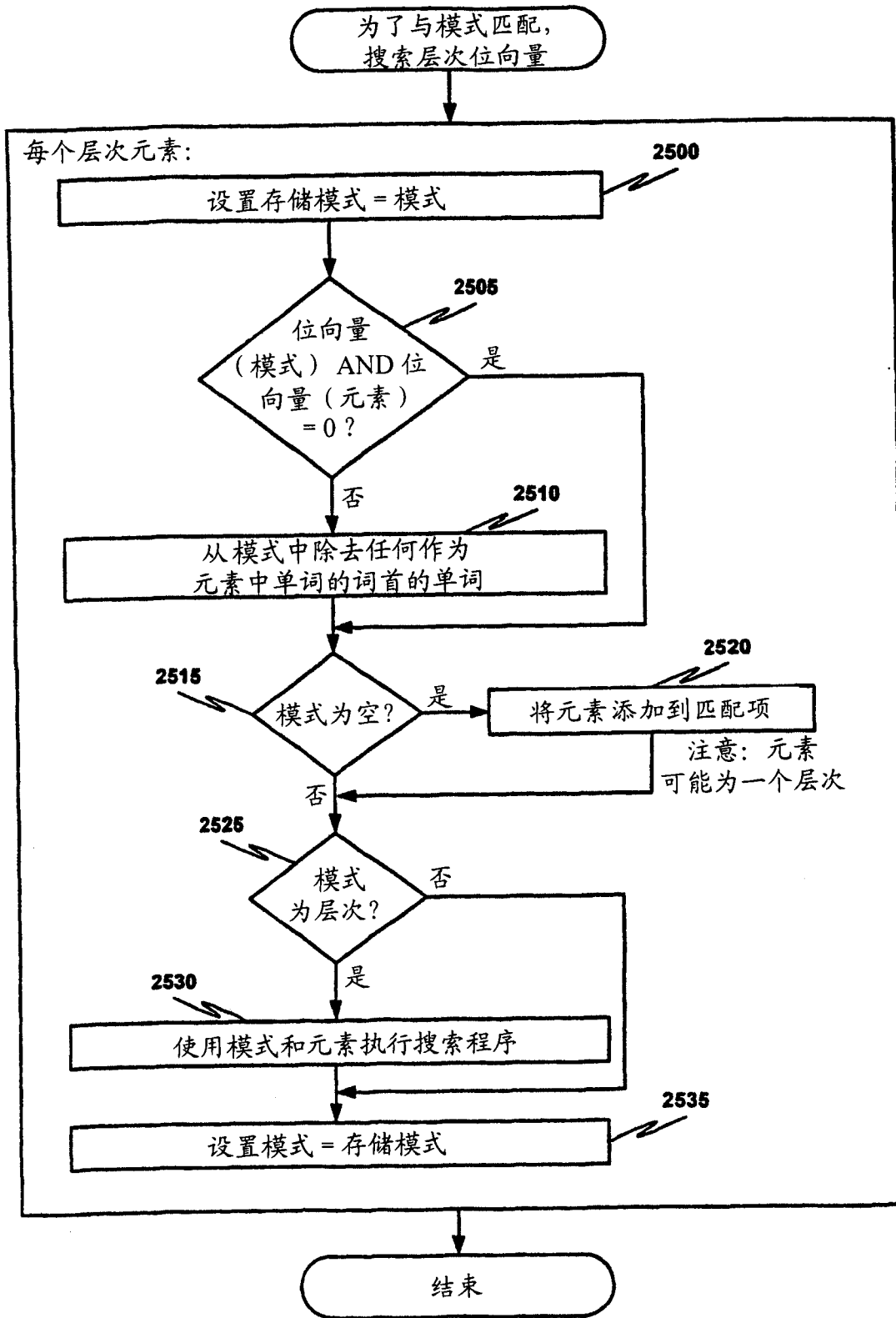


图25