





















由于 BWIICC 方法在搜索阶段仍使用 Boolean 方法的规则对节点进行扩充,只有在节点所代表的候选解非极小时才会停止对当前节点的继续扩充而返回上一层,因此,BWIICC 方法在整个求解过程中所遍历的全部节点是 Boolean 方法的子集,其正确性由 Boolean 方法保证.同时,由于非极小候选解的超集不可能是极小碰集,因此对比 Boolean 方法,使用 ICC 的 BWIICC 方法所剪枝掉的节点中不可能包含极小碰集,即,使用 ICC 后,Boolean 方法的完备性不会被破坏.

在第 3 节中,我们将会以多组实验数据对比的方式,分析说明算法 5 相较于算法 2 的优势.

### 3 实验分析

首先,将使用算法 5 给出的 BWIICC 方法与算法 2 介绍的 Boolean 方法进行比较,给出两种方法在随机生成测试用例下的实验结果.然后,会单独使用 BWIICC 方法对另外几组难度较高的测试用例进行求解,以测试这种方法对较难问题的求解能力.

实验平台如下:Windows 10 操作系统,CPU Intel Core i5-3470 3.2Ghz,8.00GB RAM,C++.

实验所用测试用例为随机生成器产生,输入参数有元素个数  $m$ ,集合簇中集合个数  $n$ ,以及元素在一个集合中出现的概率  $p$ .同一个用例中,所有元素的  $p$  均相等,每个集合包含元素的平均个数约等于  $mp$ .本文使用的实验用例共 8 小组,分为对比组和较难问题组两部分.每小组元素个数固定,对比组分别为 15,20,25,30,较难问题组分别为 35,40,45,50,各小组均包含  $p$  取值 0.05~0.94 的 19 个用例,所有用例集合簇中集合个数均为 200.本文所有实验数据均为使用相同参数下独立生成的 10 个用例进行实验,所得结果的平均值.

图 2 给出了两种方法对于对比组中所有实验用例运行时间的对比情况.

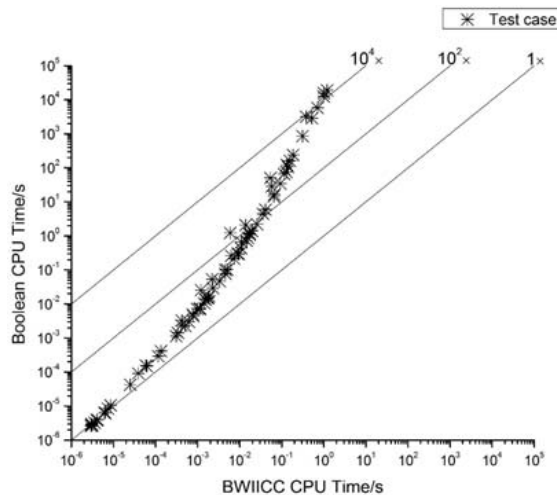


Fig.2 Experimental comparisons of BWIICC and Boolean

图 2 BWIICC 方法与 Boolean 方法实验结果比较

我们可以看出:对于耗时较少的用例,两种方法所用时间十分接近,用例点基本都落在 1 倍对比线附近;但随着用例耗时的上升,BWIICC 方法的优势逐步显现;尤其对于那些 Boolean 方法耗时在 1s 以上的用例,多数用例点已经处于  $10^2$  倍对比线上方,部分耗时较高的用例点甚至位于  $10^4$  倍对比线之上.接下来,我们将会通过实验数据的分组比较,进一步分析产生这种现象的原因.

图 3 是将两种方法的实验数据按用例元素数分组后得到的结果,其中,各子图的横坐标轴表示元素出现的概率  $p$ ,MHS 表示对应实例的极小碰集数量,与右侧纵轴相关.从图中我们可以看出:各子图中,对于大多数测试用例,产生极小碰集的数量越多,两种方法在时间效率上的差距也就越大;而对于产生极小碰集较少的用例,差距则并不明显.这主要是由于 Boolean 算法中使用的子集检测方法对极小碰集簇的大小较为敏感导致.同时,产生极小碰集的数量越多,一般也意味着用例的难度较高,算法耗时也会相应增加,这与图 2 中体现的趋势是一致的.

的.除此之外我们还可以发现,各子图中 Boolean 与 *MHS* 两条折线图拟合的并不是很好.这说明对元素个数相同的一组用例,出现概率  $p$  的不同会导致 Boolean 算法对于单一极小碰集的平均求解时间产生较大波动.

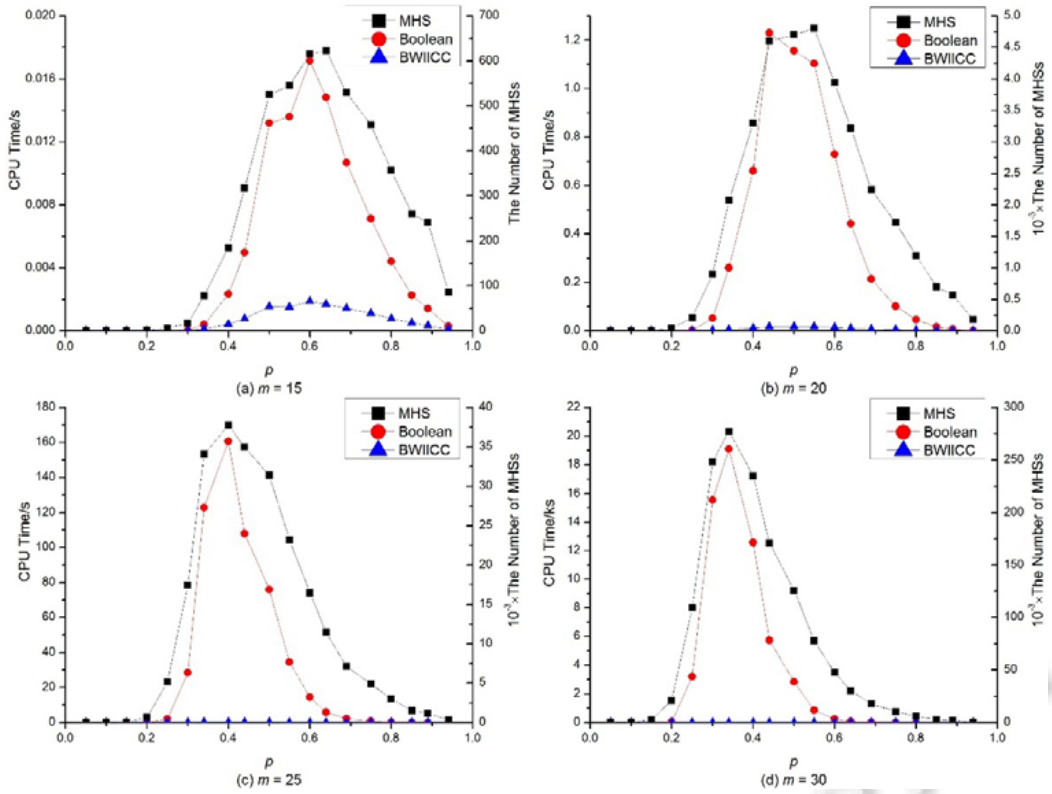


Fig.3 Experimental comparisons of the two methods based on diverse number of elements

图 3 不同元素数实例分组下 2 种方法实验结果比较

图 4 为两种方法对于各实例求解过程中访问总分支数情况的对比.分支数即是在形同图 1 的树形表示下,算法访问总节点对应的二叉树中所包含叶节点数量.其中,各子图横坐标轴表示含义与图 3 一致,纵坐标轴表示算法对于各实例求解过程中访问总分支数与该实例解中实际包含的极小碰集数的比值.以此,我们可以直观地看出使用 IICC 为 BWIICC 所带来的剪枝收益以及两种方法实际访问总分支数与理论下界(极小碰集数)的比较情况.同样对于较难的实例,IICC 所带来的剪枝效果更加明显.同时我们还能发现:虽然对于各组实例, BWIICC 访问的总分支数都明显少于 Boolean,但两者间的差距并没有图 3 中体现出的时间差那样巨大.这说明除去剪枝收益,IICC 方法在候选解极小性判定上为使用其的 BWIICC 算法带来了更为巨大的时间效率收益.

图 5 给出了 BWIICC 方法对于较难问题组测试用例的实验结果,其中,各坐标轴表示含义与图 3 中一致.

从图 5(a)中我们可以看到,算法对其中难度较高的测试用例求得的极小碰集数已达到  $1.8 \times 10^6$  个以上.而随着元素数的增加,图 5(d)中难度最高测试用例产生的极小碰集数更是超过  $6 \times 10^8$  个.但 BWIICC 方法仍可以将求解时间控制在 3000s 以内,可见, BWIICC 方法对于较高难度的问题也有着不错的求解效率.

同时,对比图 3 各子图中 Boolean 与 *MHS* 折线图的拟合情况,图 5 中 BWIICC 与 *MHS* 折线图的拟合度显然更高.这说明对于含有相同元素个数但出现概率  $p$  不同的一组用例, BWIICC 方法的单一极小碰集平均求解时间较 Boolean 方法更为稳定,随问题难度的波动幅度更小.

通过以上实验对比分析,我们得到以下结论:使用 IICC 的 BWIICC 方法相较使用子集检测的 Boolean 方法,求解效率有着显著的提升;同时,对于不同难度实例的单一极小碰集平均求解时间, BWIICC 方法也更为稳定.

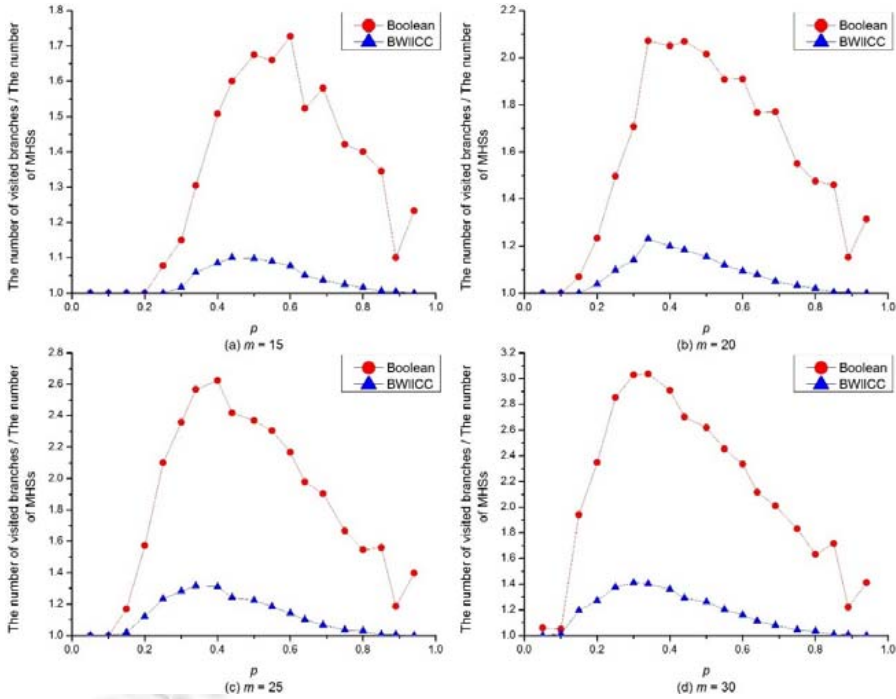


Fig.4 Experimental comparisons of the number of visited branches of the two methods  
图4 两种方法访问分支数对比

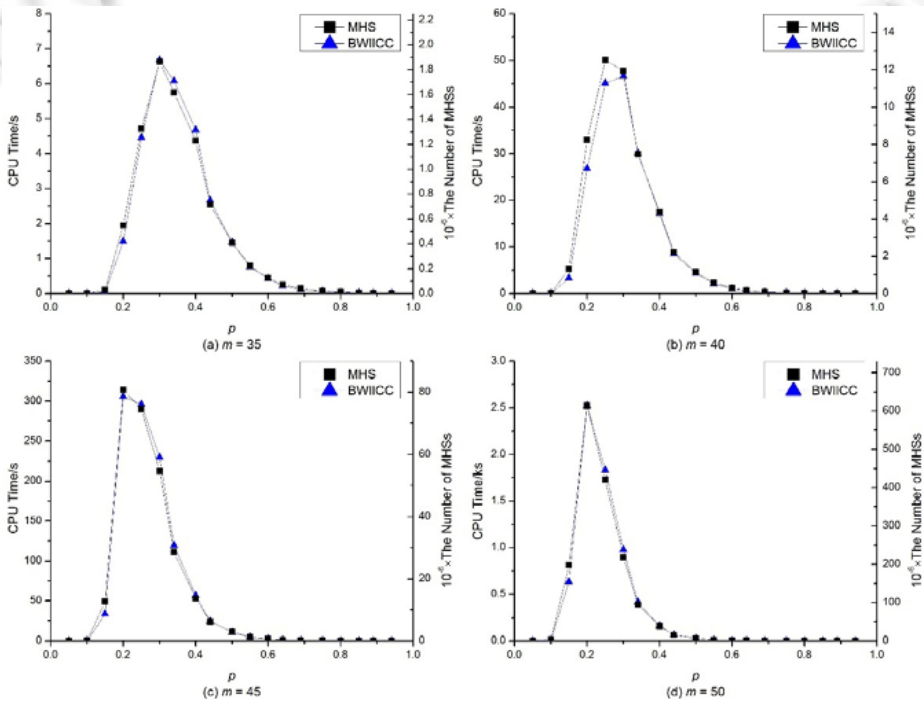


Fig.5 Experimental data of BWiCC for the difficult problems based on diverse number of elements  
图5 BWiCC 方法对于较难问题组实例的实验数据

## 4 结束语

极小碰集问题在许多重要领域都有着广泛的应用,现有的极小碰集算法主要通过使用子集检测方法来保证最终求得碰集的极小性,但这种方法效率会受到极小碰集簇大小影响,极小碰集簇中碰集越多,子集检测的耗时也就越高.对此,本文首先提出了基于元素独立覆盖度检测的碰集极小性判定方法 ICC,剥离了碰集极小性判定效率与极小碰集簇大小的相关性.随后,对增量求解过程中非极小碰集的产生原因进行了深入分析,并以此对 ICC 方法进行优化,得到 IICC 方法.该方法可以增量地对候选解的极小性给出判定结果,在提升碰集极小性判定效率的同时,还能够通过尽早发现并丢弃非极小候选解而获得额外的剪枝效果,进一步提升了极小碰集求解算法整体的执行效率.最后,将 IICC 方法与当前求解效率较高的 Boolean 算法相结合,得到 BWIICC 算法,并通过实验对比检测优化效果.实验结果表明:优化后的 BWIICC 算法相较于 Boolean 算法,整体效率提升明显,且优化效果有随问题难度上升而上升的趋势.对于对比组中难度最高的实例,效率提升可达 4 个数量级以上.同时,通过对较难问题组实例的测试,显示出 BWIICC 算法在问题难度较高时,依然可以保持不错的求解效率.此外,对于单一极小碰集的平均求解时间,BWIICC 算法也更为稳定.

**致谢** 本文作者对所有匿名审稿人的辛勤工作和宝贵意见表示真诚的感谢.

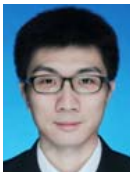
## References:

- [1] Reiter R. A theory of diagnosis from first principles. *Artificial Intelligence*, 1987,32(1):57–95.
- [2] Yu Q, Li CQ, Shen YM, Wang J. Method of solving abductive reasoning problem via hitting set. *Ruan Jian Xue Bao/Journal of Software*, 2015,26(8):1937–1945 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/4694.htm> [doi: 10.13328/j.cnki.jos.004694]
- [3] Bonet B, Helmert M. Strengthening landmark heuristics via hitting sets. In: *Proc. of the 19th European Conf. on Artificial Intelligence*. Amsterdam: IOS Press, 2010. 329–334.
- [4] Wotawa F. On the relationship between model-based debugging and program slicing. *Artificial Intelligence*, 2002,135(1):125–143.
- [5] Greiner R, Smith BA, Wilkerson RW. A correction to the algorithm in Reiter's theory of diagnosis. *Artificial Intelligence*, 1989, 41(1):79–88.
- [6] Wotawa F. A variant of Reiter's hitting-set algorithm. *Information Processing Letters*, 2001,79(1):45–51.
- [7] Jiang YF, Lin L. Computing the minimal hitting set with binary HS-tree. *Ruan Jian Xue Bao/Journal of Software*, 2002,13(12): 2267–2274 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/13/2267.htm>
- [8] Jiang YF, Lin L. The computation of hitting sets with Boolean formulas. *Chinese Journal of Computers*, 2003,26(8):919–924 (in Chinese with English abstract).
- [9] Ouyang DT, Ouyang JH, Cheng XC, Liu J. A method of computing hitting set in model-based diagnosis. *Chinese Journal of Science Instrument*, 2004,25(4):605–608 (in Chinese with English abstract).
- [10] Zhao XF, Ouyang DT. A method of combing SE-tree to compute all minimal hitting sets. *Progress in Natural Science*, 2006,16(2): 169–174.
- [11] Chen XM, Meng XF, Qiao RX. Method of computing all minimal hitting set based on BNB-HSSE. *Chinese Journal of Science Instrument*, 2010,31(1):61–67 (in Chinese with English abstract).
- [12] Zhang LM, Ouyang DT, Zeng HL. Computing the minimal hitting set based on dynamic maximum degree. *Journal of Computer Reach and Development*, 2011,48(2):209–215 (in Chinese with English abstract).
- [13] Pill I, Quaritsch T. Optimizations for the Boolean approach to computing minimal hitting sets. In: *Proc. of the 20th European Conf. on Artificial Intelligence*. Amsterdam: IOS Press, 2012. 648–653.
- [14] Pill I, Quaritsch T. RC-Tree: A variant avoiding all the redundancy in Reiter's minimal hitting set algorithm. In: *Proc. of the 2015 IEEE Int'l Symp. on Software Reliability Engineering Workshops*. 2015. 78–84.
- [15] Wang YY, Ouyang DT, Zhang LM, Zhang YG. A method of computing minimal hitting sets using CSP. *Journal of Computer Reach and Development*, 2015,52(3):588–595 (in Chinese with English abstract).
- [16] Lin L, Jiang YF. Computing minimal hitting sets with genetic algorithm. In: *Proc. of the 13th Int'l Workshop on Principles of Diagnosis*. 2002. 77–80.
- [17] Cincotti A, Cutello V, Pappalardo F. An ant-algorithm for the weighted minimum hitting set problem. In: *Proc. of the 2003 IEEE Symp. on Swarm Intelligence*. Piscataway: IEEE, 2003. 1–5.

- [18] Huang J, Chen L, Zou P. Computing minimal diagnosis by compounded genetic and simulated annealing algorithm. Ruan Jian Xue Bao/Journal of Software, 2004,15(9):1345–1350 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/15/1345.htm>
- [19] Abreu R, Gemund AJCV. A low-cost approximate minimal hitting set algorithm and its application to model-based diagnosis. In: Proc. of the 8th Symp. on Abstraction, Reformulation, and Approximation, Vol.9. Menlo Park: AAAI, 2009. 2–9.
- [20] Zhou G, Feng WQ, Jiang BF, *et al.* Computing minimal hitting set based on immune genetic algorithm. Int'l Journal of Modelling, Identification and Control, 2014,21(1):93–100.
- [21] Jannach D, Schmitz T, Shchekotykhin K. Parallelized hitting set computation for model-based diagnosis. In: Proc. of the 29th AAAI Conf. on Artificial Intelligence. Menlo Park: AAAI, 2015. 1503–1510.
- [22] Zhao XF, Ouyang DT. Deriving all minimal hitting sets based on join relation. IEEE Trans. on Systems, Man, and Cybernetics: Systems, 2015,45(7):1063–1076.
- [23] Pill I, Quaritsch T, Wotawa F. From conflicts to diagnoses: An empirical evaluation of minimal hitting set algorithms. In: Proc. of the 22nd Int'l Workshop on the Principles of Diagnosis. 2011. 203–210.
- [24] Han B, Lee SJ. Deriving minimal conflict sets by CS-tree with mark set in diagnosis from first principles. IEEE Trans. on Systems, Man, and Cybernetics: Cybernetics, 1999,29(2):281–286.

#### 附中文参考文献:

- [2] 余泉,李承乾,申宇铭,王驹.溯因推理问题的碰集求解方法.软件学报,2015,26(8):1937–1945. <http://www.jos.org.cn/1000-9825/4694.htm> [doi: 10.13328/j.cnki.jos.004694]
- [7] 姜云飞,林笠.用对分 HS-树计算最小碰集.软件学报,2002,13(12):2267–2274. <http://www.jos.org.cn/1000-9825/13/2267.htm>
- [8] 姜云飞,林笠.用布尔代数方法计算最小碰集.计算机学报,2003,26(8):919–924.
- [9] 欧阳丹彤,欧阳继红,程晓春,刘杰.基于模型诊断中计算碰集的方法.仪器仪表学报,2004,25(4):605–608.
- [11] 陈晓梅,孟晓风,乔仁晓.基于 BNB-HSSE 计算全体碰集的方法.仪器仪表学报,2010,31(1):61–67.
- [12] 张立明,欧阳丹彤,曾海林.基于动态极大度的极小碰集求解方法.计算机研究与发展,2011,48(2):209–215.
- [15] 王艺源,欧阳丹彤,张立明,张永刚.利用 CSP 求解极小碰集的方法.计算机研究与发展,2015,52(3):588–595.
- [18] 黄杰,陈琳,邹鹏.一种求解极小诊断的遗传模拟退火算法.软件学报,2004,15(9):1345–1350. <http://www.jos.org.cn/1000-9825/15/1345.htm>



刘思光(1988—),男,吉林省吉林市人,硕士,主要研究领域为基于模型诊断,极小碰集求解.



张立明(1980—),男,博士,高级工程师,CCF 专业会员,主要研究领域为基于模型诊断,SAT.



欧阳丹彤(1968—),女,博士,教授,博士生导师,CCF 高级会员,主要研究领域为基于模型诊断.