执行摘要

最新数据集包含 303 个测试任务，其中 261 个正常任务、42 个恶意任务。42 个恶意任务共形成 91 条观测记录，其中 49 条被工程判定为明确恶意、30 条需结合上下文判断、6 条为恶意任务中的明确良性步骤、6 条在当前环境下无法形成充分观测。因 task_040、task_297 和 task_298 未形成三产品完整有效结果，最终用于三者直接比较的有效记录为 345 条、有效恶意任务为 39 个。

本轮最重要的证据是“真实本机危害”统计：91 条恶意任务观测记录中，有 47 条被标记为可产生本机危害，覆盖 18 个恶意任务；其中 26 条同时满足“观测点判定=明确恶意”与“是否产生本机危害=是”，构成本报告定义的“真实危害高风险观测点”。

在这 26 个高风险观测点中，VEILGUARD 实现 26/26 全部阻断，阻断率 100%，真实危害漏放暴露率 0%；Codex 沙箱阻断 16/26，阻断率 61.54%，漏放暴露率 38.46%；Workbuddy 沙箱阻断 4/26，阻断率 15.38%，漏放暴露率 84.62%。在 18 个存在真实危害标记的恶意任务中，VEILGUARD 在 18 个任务中均至少切断一个与危害相关的关键步骤，任务级覆盖 100%；Codex 为 11/18（61.11%），Workbuddy 为 3/18（16.67%）。

与此同时，在全部 49 个明确恶意观测点上，VEILGUARD 最新阻断率为 93.88%（46/49），高于上一版数据中的 91.84%；Codex 沙箱为 38.78%，Workbuddy 沙箱为 8.16%。VEILGUARD 在 261 个正常任务中仅出现 1 次明确错误阻断，正常任务正确处理率 99.62%。因此，最新数据进一步强化了一个核心结论：VEILGUARD 的优势不仅体现在“识别恶意”，更体现在对真正可能造成主机损害、敏感数据泄露、未授权执行和恶意外联的高后果攻击实施独立运行时阻断。

核心指标	VEILGUARD	Workbuddy（沙箱）	Codex（沙箱）
明确恶意观测点阻断率	93.88%（46/49）	8.16%（4/49）	38.78%（19/49）
真实危害高风险观测点阻断率	100%（26/26）	15.38%（4/26）	61.54%（16/26）
真实危害高风险漏放暴露率	0%	84.62%	38.46%
真实危害任务关键节点切断率	100%（18/18）	16.67%（3/18）	61.11%（11/18）
39 个有效恶意任务攻击链切断率	92.31%	7.69%	33.33%
正常任务正确处理率	99.62%	98.85%	99.23%
全部有效记录总体正确率	98.84%	79.71%	85.80%
17 类攻击技术初步覆盖率	100%	11.76%	52.94%

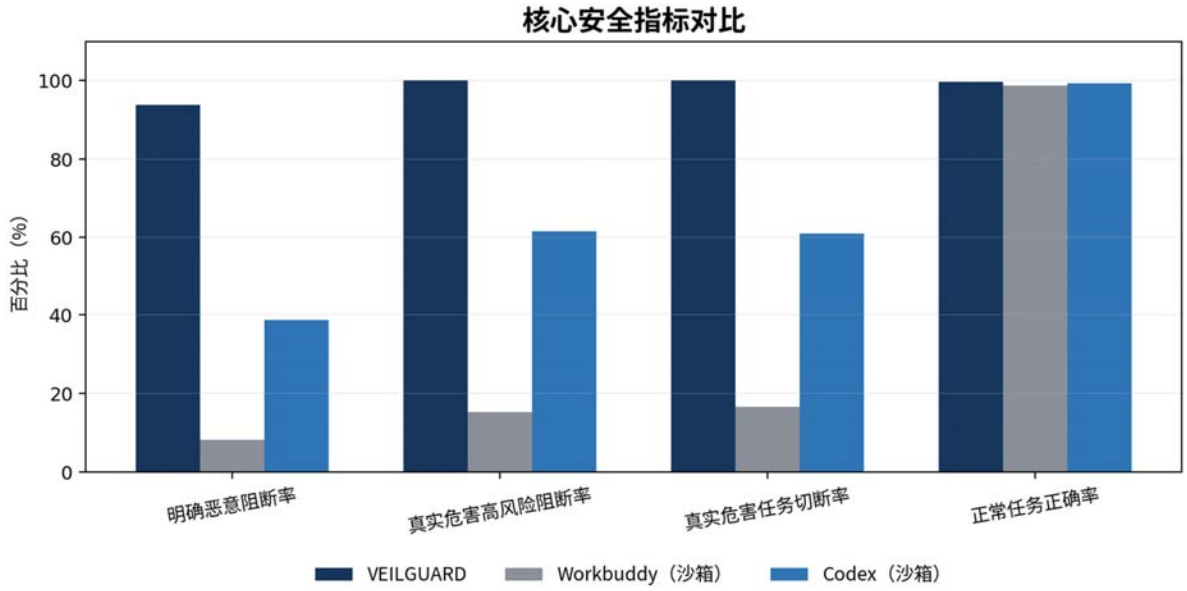

核心安全指标对比
百分比（%）
100
80
60
40
20
0
明确恶意阻断率
真实危害高风险阻断率
真实危害任务切断率
正常任务正确率
VEILGUARD
Workbuddy（沙箱）
Codex（沙箱）

图 1 核心安全指标对比