

注：作弊等级分六级（0~5，0表示没有过滤，5表示作弊嫌疑最大，一般>=4即可定义为高危）

数据文件格式为 SequenceFile, key/value 均为 Text

天级别数据保留一个月，小时级别全量 actionid 数据保留 2 天，小时级别其它数据保留 7 天

级别	数据路径举例	数据说明
天级别	/user/jianshan.he/antispam-rule-engine/day_result/20151109/ FILTERED_ACTIONID -r- key: actionid value: 空	该天过滤的全部 actionid
	/user/jianshan.he/antispam-rule-engine/day_result/20151109/ IP -r- key: ip value: 作弊等级^A 过滤比^A 过滤数^A 总数^A 过滤详情....	该天 IP 的过滤汇总
	/user/jianshan.he/antispam-rule-engine/day_result/20151109/ PYID -r- key: pyid value: 作弊等级^A 过滤比^A 过滤数^A 总数^A 过滤详情....	该天 pyid 的过滤汇总
	/user/jianshan.he/antispam-rule-engine/day_result/20151109/ DOMAIN -r- key: domain value: 作弊等级^A 过滤比^A 过滤数^A 总数^A 过滤详情....	该天完整域名的过滤汇总
	/user/jianshan.he/antispam-rule-engine/day_result/20151109/ TLDOMAIN -r- key: 顶级域名 value: 作弊等级^A 过滤比^A 过滤数^A 总数^A 过滤详情....	该天顶级域名的过滤汇总
	/user/jianshan.he/antispam-rule-engine/day_result/20151109/ PLATFORM_DOMAIN -r- key: platform^Adomain value: 作弊等级^A 过滤比^A 过滤数^A 总数^A 过滤详情....	该天平台加完整域名的过滤汇总
	/user/jianshan.he/antispam-rule-engine/day_result/20151109/ ADUNIT -r- key: platform^A 顶级域名^Aadunitid^AadWidth^AadHeight^AadLocation value: 作弊等级^A 过滤比^A 过滤数^A 总数^A 过滤详情....	该天点位的过滤汇总
小时级别	/user/jianshan.he/antispam-rule-engine/hour_result/20151109/00/all_actions/part- key: actionid value: 过滤原因 1^A 过滤原因 2^A....^Bip^Apyid^Adomain^Aplatform^Aadunitid^AadWidth^AadHeight^AadLocation	该小时全量 actionid
	/user/jianshan.he/antispam-rule-engine/hour_result/20151109/00/filters/ FILTERED_ACTIONID -r- key: actionid value: 空	该小时过滤的 actionid
	/user/jianshan.he/antispam-rule-engine/hour_result/20151109/00/filters/ IP -r- key: ip value: 作弊等级^A 过滤比^A 过滤数^A 总数^A 过滤详情....	该小时 IP 的过滤汇总
	/user/jianshan.he/antispam-rule-engine/hour_result/20151109/00/filters/ PYID -r- key: pyid value: 作弊等级^A 过滤比^A 过滤数^A 总数^A 过滤详情....	该小时 pyid 的过滤汇总
	/user/jianshan.he/antispam-rule-engine/hour_result/20151109/00/filters/ DOMAIN -r- key: domain value: 作弊等级^A 过滤比^A 过滤数^A 总数^A 过滤详情....	该小时完整域名的过滤汇总
	/user/jianshan.he/antispam-rule-engine/hour_result/20151109/00/filters/ TLDOMAIN -r- key: 顶级域名 value: 作弊等级^A 过滤比^A 过滤数^A 总数^A 过滤详情....	该小时顶级域名的过滤汇总
	/user/jianshan.he/antispam-rule-engine/hour_result/20151109/00/filters/ PLATFORM_DOMAIN -r- key: platform^Adomain value: 作弊等级^A 过滤比^A 过滤数^A 总数^A 过滤详情....	该小时平台加完整域名的过滤汇总
	/user/jianshan.he/antispam-rule-engine/hour_result/20151109/00/filters/ ADUNIT -r- key: platform^A 顶级域名^Aadunitid^AadWidth^AadHeight^AadLocation value: 作弊等级^A 过滤比^A 过滤数^A 总数^A 过滤详情....	该小时点位的过滤汇总