

Malware Detection Dataset

Based on the characteristics of the observations, the dataset was created in a Unix / Lunix-based virtual machine for classification purposes, which are harmless with malware software for Android devices. The data set consists of 100,000 observation data and 35 features. Below is a table of specifications and descriptions.

| Features | Description |
|-------------------|---|
| hash | APK/ SHA256 file name |
| milisecond | time |
| classification | malware/beign |
| state | flag of unrunable/runnable/stopped tasks. |
| usage_counter | task structure usage counter |
| prio | keeps the dynamic priority of a process |
| static_prio | static priority of a process |
| normal_prio | priority without taking RT-inheritance into account |
| policy | planning policy of the process |
| vm_pgoff | the offset of the area in the file, in pages. |
| vm_truncate_count | used to mark a vma as now dealt with |
| task_size | size of current task. |
| cached_hole_size | size of free address space hole. |
| free_area_cache | first address space hole |
| mm_users | address space users |
| map_count | number of memory areas |
| hiwater_rss | peak of resident set size |
| total_vm | total number of pages |
| shared_vm | number of shared pages. |
| exec_vm | number of executable pages. |
| reserved_vm | number of reserved pages. |
| nr_ptes | number of page table entries |
| end_data | end address of code component |
| last_interval | last interval time before thrashing |
| nvcs | number of volunteer context switches. |
| nivcs | number of in-volunteer context switches |
| minflt | minor page faults |
| majflt | major page faults |
| fs_excl_counter | it holds file system exclusive resources. |
| lock | the read-write synchronization lock used for file system access |
| utime | user time |
| stime | system time |
| gtime | guest time |
| cgroup | cumulative group time. Cumulative resource counter |
| signal_nvcs | used as cumulative resource counter. |