

CLAIMS

1. ~~_____~~ A computer implemented method for maintaining data processing system security, the method implemented by one or more computer processors comprising:

receiving a plurality of streams, each corresponding to a respective data file, into a data store;

for each of the corresponding respective data files:

calculating a respective checksum using at least part of a corresponding respective stream of the plurality of streams,

selecting respective corresponding metadata from a metadata storage unit, ~~by one or more computer processors, metadata associated by matching the respective calculated checksum with at least one part of a data file;~~ a checksum stored in correspondence with the metadata, and

retrieving respective one or more virus indicators from the selected respective corresponding metadata;

generating, ~~a prioritized list including the each corresponding respective data files by the one or more computer processors, a rating the priority for the at least one part of the data file according to the metadata; and~~

conducting, ~~each corresponding respective data files in order of a respective likelihood, determined by the respective one or more computer processors, a scan~~ retrieved virus indicators, that the ~~each of the data file according to the priority~~ files contains malware;

wherein ~~the metadata comprises~~ conducting a scan of a first data file having a highest priority in the prioritized list before conducting a scan of another data file in the prioritized list;

after the conducting, receiving a new stream corresponding to a new data file;

retrieving ~~new~~ one or more virus indicators ~~corresponding to the new data file from newly selected corresponding metadata from the metadata storage unit;~~

re-generating the prioritized list, including the each corresponding respective data file that have not been scanned and the new data file, by rating the priority of the each corresponding respective data files that have not been scanned and the new data file in order of a respective likelihood, determined by the respective one or more retrieved virus indicator and the new one or more virus indicators, that the each of the data files and the new data file contains malware;

scanning each file in the regenerated prioritized list in order, beginning with the highest prioritized file in the re-generated prioritized list; and

blocking user access to each file in the re-generated prioritized list until after the each file is scanned.

2. ———The computer implemented method according to claim 1, wherein each respective checksum is continuously calculated in parallel as each corresponding respective data file streams into the data store.

3. The computer implemented method according to claim 1, further comprising: updating, by the one or more computer processors, each respective corresponding metadata of each respective corresponding data file according to the scan of the each respective corresponding data file; and providing, by the one or more computer processors, each respective corresponding data file for further processing.

4. The computer implemented method according to claim 1, wherein the metadata is associated with a checksum continuously calculated in parallel as the data file streams into a data store.

~~3. ——— The computer implemented method according to either of the preceding claims, further comprising:~~

~~updating, by the one or more computer processors, the metadata of the data file according to the scan; and~~

~~providing, by the one or more computer processors, the data file for further processing.~~

~~4. ——— The computer implemented method according to claim 1, wherein the priority prioritised list is generated according to a correlation of data file metadata.~~

5. ———The computer implemented method according to claim 1, further comprising presenting, by the one or more computer processors, at least one of the respective data filefiles to a cleansing routine.

6. ———The computer implemented method according to any of the preceding claims, further comprising:

~~matching, by the one or more computer processors, a checksum for a part of the data file with a checksum in the metadata storage unit; and~~

~~triggering, by the one or more computer processors, a data cleansing routine in response to the matching.~~

~~7. The computer implemented method according to any of the preceding claims~~claim 1, further comprising updating, by the one or more computer processors, ~~checksum~~one of the checksums stored in correspondence with the metadata associated with a part of the data file.

~~8. 7. A computer implemented method~~program product for maintaining system security, the ~~method~~computer program product comprising:

~~receiving, by one or more computer processors, an updated virus definition;~~

~~obtaining, by the one or more computer processors, metadata associated with at least one~~readable storage devices including non-transitory computer readable memory including stored program instructions thereon, the stored program instructions, when executed, causing the execution of steps comprising:

receiving a plurality of streams, each corresponding to a respective data file-, into a data store;

for each of the corresponding respective data files;

calculating a respective checksum using at least part of a corresponding respective stream of the plurality of streams,

selecting respective corresponding metadata from a metadata storage unit; by matching the respective calculated checksum with a checksum stored in correspondence with the metadata, and

retrieving respective one or more virus indicators from the selected respective corresponding metadata;

generating a prioritized list including the each corresponding respective data files by rating the priority of the each corresponding respective data files in order of a respective likelihood, determined by the respective one or more retrieved virus indicators, that the each of the data files contains malware;

conducting a scan of a first data file having a highest priority in the prioritized list before conducting a scan of another data file in the prioritized list;

after the conducting, receiving a new stream corresponding to a new data file;

retrieving new one or more virus indicators corresponding to the new data file from newly selected corresponding metadata from the metadata storage unit;

re-generating the prioritized list, including the each corresponding respective data files that have not been scanned and the new data file, by rating the priority of the each corresponding respective data files that have not been scanned and the new data file in order of a respective likelihood, determined by

the respective one or more retrieved virus indicators and the new one or more virus indicators, that the each of the data files and the new data file contains malware;

scanning each file in the re-generated prioritized list in order, beginning with the highest prioritized file in the re-generated prioritized list; and

blocking user access to each file in the re-generated prioritized list until after the each file is scanned.

8. The computer program product according to claim 7, wherein each respective checksum is continuously calculated in parallel as each corresponding respective data file streams into the data store.

9. The computer program product according to claim 7, further comprising: program instructions to update each respective corresponding metadata of each respective corresponding data file according to the scan of the each respective corresponding data file; and provide each respective corresponding data file for further processing.

10. The computer program product according to claim 7, wherein the prioritized list is generated according to a correlation.

11. The computer program product according to claim 7, further comprising program instructions to present at least one of the respective data files to a cleansing routine.

12. The computer program product according to claim 7, further comprising program instructions to update one of the checksums stored in correspondence with the metadata.

13. ~~generating, by~~A computer system for maintaining system security, the computer system comprising: one or ~~more~~ computer processors, ~~a priority for the at least one data file according to the metadata;~~ and ~~updated virus definition;~~ ———

~~selecting, one or more computer readable storage devices including non-transitory computer readable memory including, stored program instructions thereon for execution~~ by the one or more computer processors, ~~a the stored program instructions comprising program instructions to:~~

receive a plurality of streams, each corresponding to a respective data file, into a data store;
for each of the corresponding respective data files;

calculate a respective checksum using at least part of a corresponding respective stream of the plurality of streams,

select respective corresponding metadata from a metadata storage unit by matching the respective calculated checksum with a checksum stored in correspondence with the metadata, and

retrieve respective one or more virus indicators from the selected respective corresponding metadata;

generate a prioritized list including the each corresponding respective data files by rating the priority of the each corresponding respective data files in order of a respective likelihood, determined by the respective one or more retrieved virus indicators, that the each of the data files contains malware;

conduct a scan of a first data file according to the having a highest priority;

——— in the prioritized list before conducting, by the one or more computer processors, a scan of the another data file in the prioritized list;

updating, by the one or more computer processors, the metadata of the data file according to the scan; and

providing, by the one or more computer processors, the data file for further processing.

9. ——— The computer implemented method according to claim 8, wherein the metadata is associated with a checksum continuously calculated in parallel as the data file streams into a data store.

10. ——— A computer implemented method for maintaining system security, the method comprising:

obtaining, by one or more computer processors, metadata associated with at least one after the conducting, receive a new stream corresponding to a new data file;

retrieve new one more virus indicators corresponding to the new data file from a newly selected corresponding metadata from the metadata storage unit;

generating, by the one or more computer processors, a priority for the at least one data file according to the metadata; ———

selecting, by the one or more computer processors, a data re-generate the prioritized list, including the each corresponding respective data files that have not been scanned and the new data file, by rating the priority of the each corresponding respective data files that have not been scanned and the new data file in order of a respective likelihood, determined by the respective one or more retrieved virus indicators and the new one or more virus indicate, that the each of the data files and the new data file contains malware;

scan each file according to the priority;
conducting, by in the one or more computer processors, a scan of re-generated prioritized list in
order, beginning with the data highest prioritized file;
determining, by in the one or more computer processors, a common vulnerabilities and
exposures score for the data file; re-generated prioritized list; and
updating, by the one or more block user access to each file in the re-generated prioritized list
until after the each file is scanned.

14. The computer processors, the metadata of the data file according to the scan; and
providing, by the one or more computer processors, the score to a system user for evaluation.

11. The computer implemented method system according to claim 10 13, wherein the metadata is
associated with a each respective checksum is continuously calculated in parallel as the each
corresponding respective data file streams into a the data store.

12. A15. The computer system for maintaining system security, the computer system according to
claim 13, further comprising:

one or more computer processors;
one or more computer readable storage devices; and
stored program instructions on the one or more computer readable storage devices for execution
by the one or more computer processors, the stored program instructions comprising:
program instructions to receive from a to update each respective corresponding metadata storage
unit metadata associated with at least one data file;
program instructions to generate a priority for the at least one of each respective corresponding
data file according to the metadata; and
program instructions to conduct a scan of the each respective corresponding data file; and provide each
respective corresponding data file according to the priority; for further processing.
wherein the metadata comprises one or more virus indicators.

13.

16. The computer system according to claim 12 13, wherein the metadata is associated with a
checksum continuously calculated in parallel as the data file streams into a data store prioritized list is
generated according to a correlation.

~~14.—17.~~ The computer system ~~according~~acceding to ~~either of claims 12 or~~claim 13, further comprising:

~~program instructions to update the metadata of the data file according to the scan; and
program instructions to provide the data file for further processing.~~

~~15.—The computer system according to any of claims 12 to 14, wherein the priority is generated according to a correlation of data file metadata.~~

~~16.—The computer system according to any of claims 12 to 15, further comprising~~ program instructions to present at least one of the respective data ~~file~~files to a cleansing routine.

18.

~~17.—The computer~~compute system according to any of claims 12 to 16, further comprising:
~~program instructions to match, by the one or more computer processors, a checksum for a part of the data file with a checksum in the metadata storage unit; and
program instructions to trigger a data cleansing routine in response to matching the checksum.~~

~~18.—The computer system according to any of claims 12 to 17~~claim 13, further comprising program instructions to update ~~checksum metadata associated with a part~~one of the data file.

~~19.—A computer program product for maintaining system security, the computer program product comprising:~~

~~a computer readable storage medium readable by a processing circuit and storing instructions for execution by the processing circuit for performing a method according to any of claims 1 to 11.~~

~~20.—A computer program~~checksums stored on a computer readable medium and loadable into the internal memory of a digital computer, comprising software code portions, when said program is run on a computer, for performing the method of any of claims 1 to 11.

~~-in correspondence with the metadata.~~