Claims

1. A computer-implemented method of detecting potential cybersecurity threats from collected data pertaining to a monitored network, the collected data comprising at least one of network data and endpoint data, the method comprising:

structuring the collected data as at least one data matrix, each row of the data matrix being a datapoint and each column corresponding to a feature;

identifying at least one of the datapoints as anomalous, thereby detecting a potential cybersecurity threat, wherein the datapoint is identified as anomalous based on an anomaly score computed as:

a sum of squared components of a corresponding row of a residuals matrix, or

a sum of squared components of a second-pass coordinate vector for the datapoint; and

extracting causal information about the anomalous datapoint based on an angular relationship between ~~a~~ the second-pass coordinate vector of the anomalous datapoint and a second-pass coordinate vector of at least one of the features,

the second-pass coordinate vectors determined by applying a second-pass singular value decomposition (SVD) to ~~a~~ the residuals matrix,

the residuals matrix computed between the data matrix and an approximation of the data matrix, by applying a first-pass truncated SVD to the data matrix.

~~2. The method of claim 1, wherein the datapoint is identified as anomalous based on:~~

~~a row of the residuals matrix corresponding to the datapoint, or~~

~~the second-pass coordinate vector for the datapoint.~~

~~3. The method of claim 2, wherein the datapoint is identified as anomalous based on an anomaly score computed as:~~

~~a sum of squared components of the corresponding row of the residuals matrix, or~~

~~a sum of squared components of the second-pass coordinate vector for the datapoint.~~

~~4.~~2. The method of claim 1~~, 2 or 3~~, wherein the causal information is extracted based on (i) the angular relationship and (ii) magnitude information about the second-pass coordinate vector of the feature.

~~5.~~3. The method of claim 1~~, 2, 3 or 4~~, comprising applying an unsupervised classification process to an initial data matrix, the initial data matrix comprising rows corresponding to the datapoints, to classify each row of the initial data matrix in relation to a set of classes.

~~6.~~4. The method of claim ~~5~~ 3, wherein the unsupervised classification process comprises:

applying an initial SVD to the initial data matrix, to determine an initial coordinate vector for each row of the initial data matrix, and

applying an unsupervised clustering algorithm to the initial coordinate vectors to determine a plurality of clusters of those initial coordinate vectors.

~~7.~~5. The method of ~~any preceding~~ claim 3, wherein the data matrix to which the first-pass SVD is applied has (i) at least one class column indicating classes to which the datapoints are assigned and (ii) feature columns corresponding to other features.

~~8.~~6. The method of claim ~~7 when dependent on claim 5 or 6~~ 5, comprising applying an unsupervised classification process to an initial data matrix, the initial data matrix comprising rows corresponding to the datapoints, to classify each row of the initial data matrix in relation to a set of classes, wherein the initial data matrix to which the unsupervised classification process is applied comprises said feature columns, ~~whereby~~ wherein the data points are classified based on said other features to determine the at least one class column for applying the first-pass SVD.

~~9.~~7. The method of claim ~~8~~ 6, wherein the datapoints are re-clustered by applying an unsupervised clustering algorithm to first-pass coordinate vectors thereof, the first-pass coordinate vectors determined in the first-pass SVD.

~~10.~~8. The method of claim ~~5 or 6~~ 3, wherein the data is partitioned into multiple partitions based on the results of the unsupervised classification, wherein the data matrix to which the first-pass SVD is applied corresponds to a single partition of the multiple partitions, whereby the residuals matrix is specific to that single partition.

~~11.~~9. The method of claim ~~10~~ 8, wherein the collected data is structured into multiple data matrices, each corresponding to a different one of the partitions, and the identifying and extracting steps are applied separately to each of the multiple data matrices.

~~12. The method of claim 10 or 11 when dependent on claim 6, wherein each partition corresponds to a single cluster of the plurality of clusters.~~

~~13. The method of any preceding claim, wherein the second-pass SVD is truncated.~~

~~14. The method of any of claims 1 to 12, wherein the second-pass SVD is non-truncated.~~

~~15.~~10. The method of ~~any preceding~~ claim 1, comprising the step of controlling a user interface to provide, to an analyst, the causal information about the anomalous datapoint.

~~16.~~11. The method of claim ~~15 when dependent on claim 3~~10,

wherein the causal information is provided in response to determining that an escalation condition is met, the escalation condition dependent on the anomaly score.

~~17.~~12. The method of claim ~~16~~ 11, wherein the collected data is used to create and populate a set of case records, each case record associated with a threat score, wherein at least one of the cases pertains to the anomalous datapoint and the threat score of that case is dependent on the anomaly score, wherein the case is rendered available via the user interface when its threat score meets the escalation condition and the causal information is provided as part of the case record.

~~18.~~13. The method of ~~any preceding~~ claim 1, wherein each datapoint corresponds to a network endpoint, the method being applied to identify at least one anomalous network endpoint.

~~19. The method of claim 18, wherein the features pertain to multiple processes running on each endpoint.~~

~~20.~~14. The method of ~~any of claims 1 to 17~~ claim 1, wherein the data matrix corresponds to a single process, each datapoint corresponding to an instance of that single process, the method being applied to identity at least one anomalous instance of that process.

~~21. The method of claim 20, wherein the collected data is partitioned according to process, to obtain multiple partitions each corresponding to a different single process, wherein the identifying and extracting steps are applied separately for each partition.~~

~~22. The method of claim 20 or 21, wherein the features pertain to network addresses accessed by the process instances.~~

~~23.~~15. The method of ~~any of claims 1 to 17~~ claim 1, wherein each datapoint corresponds to a network address, the method being applied to identify at least one anomalous network address.

~~24.~~16. A ~~computer-implemented~~ computer system comprising one or more hardware computers programmed or otherwise-configured to carry out a method of detecting potential cybersecurity threats from collected data pertaining to a monitored network, the collected data comprising at least one of network data and endpoint data, the method comprising:

structuring the collected data as ~~an initial~~ at least one data matrix, each row of the data matrix being a datapoint and each column corresponding to a feature;

~~applying an unsupervised classification process to the initial data matrix to classify each datapoint in relation to a set of classes;~~

~~re-structuring the collected data as at least one data matrix for use in anomaly detection, the at least one data matrix determined using the results of the unsupervised classification process for incorporating class information into the anomaly detection; and~~

~~applying anomaly detection to the at least one data matrix, thereby identifying at least one of the datapoints as anomalous using a residuals matrix, the residuals matrix computed between the data matrix and an approximation of the data matrix, by applying a truncated singular value decomposition (SVD) to the data matrix.~~

identifying at least one of the datapoints as anomalous, thereby detecting a potential cybersecurity threat, wherein the datapoint is identified as anomalous based on an anomaly score computed as:

a sum of squared components of a corresponding row of a residuals matrix, or a sum of squared components of a second-pass coordinate vector for the datapoint; and

extracting causal information about the anomalous datapoint based on an angular relationship between the second-pass coordinate vector of the anomalous datapoint and a second-pass coordinate vector of at least one of the features,

the second-pass coordinate vectors determined by applying a second-pass singular value decomposition (SVD) to the residuals matrix,

the residuals matrix computed between the data matrix and an approximation of the data matrix, by applying a first-pass truncated SVD to the data matrix.

~~25. The method of claim 24, wherein the datapoint is identified as anomalous based on a row of the residuals matrix corresponding to the datapoint.~~

~~26. The method of claim 24 or 25, comprising the step of extracting causal information about the anomalous datapoint based on an angular relationship between a second-pass coordinate vector of the anomalous datapoint and a second-pass coordinate vector of at least one of the features, the second-pass coordinate vectors determined by applying a second-pass SVD to the residuals matrix.~~

~~27. The method of claim 26, wherein the datapoint is identified as anomalous based on the second-pass coordinate vector of the datapoint.~~

~~28. The method of any of claims 24 to 27, wherein the results of the unsupervised classification process are used to determine an augmented data matrix with one or more classification features, the anomaly detection being applied to the augmented data matrix.~~

~~29. The method of any of claims 24 to 27, wherein the collected data is partitioned into multiple partitions based on the results of the unsupervised classification process, wherein the data matrix to which the anomaly detection is applied corresponds to a single partition of the multiple partitions, whereby the residuals matrix is specific to that single partition.~~

~~30. The method of claim 29, wherein the collected data is structured into multiple data matrices, each corresponding to a different one of the partitions, and the identifying and extracting steps are applied separately to each of the multiple data matrices.~~

~~31. The method of any of claims 24 to 30, wherein the unsupervised classification process comprises:~~

~~applying an initial SVD to the initial data matrix, to determine an initial coordinate vector for each row of the initial data matrix, and~~

~~applying an unsupervised clustering algorithm to the initial coordinate vectors to determine a plurality of clusters of those initial coordinate vectors.~~

~~32.~~17. A ~~computer-implemented~~ non-transitory computer-readable medium comprising program instructions for programming a computer or a set of computers to carry out an anomaly detection method, the method comprising:

determining a data matrix from a set of collected data, each row of the data matrix being a datapoint and each column corresponding to a feature;

identifying at least one of the datapoints as anomalous, wherein the datapoint is identified as anomalous based on an anomaly score computed as:

a sum of squared components of a corresponding row of a residuals matrix, or a sum of squared components of a second-pass coordinate vector for the datapoint; and

extracting causal information about the anomalous datapoint based on an angular relationship between ~~a~~ the second-pass coordinate vector of the anomalous datapoint and a second-pass coordinate vector of at least one of the features,

the second-pass coordinate vectors determined by applying a second-pass singular value decomposition (SVD) to ~~a~~ the residuals matrix,

the residuals matrix computed between the data matrix and an approximation of the data matrix, by applying a first-pass truncated SVD to the data matrix.

~~33. A computer-implemented anomaly detection method, the method comprising:~~

~~determining an initial data matrix from a set of collected data, each row of the initial data matrix being a datapoint and each column corresponding to a feature;~~

~~applying an unsupervised classification process to the initial data matrix to classify each datapoint in relation to a set of classes;~~

~~re-structuring the collected data as at least one data matrix for use in anomaly detection, the at least one data matrix determined using the results of the unsupervised classification process for incorporating class information into the anomaly detection; and~~

applying anomaly detection to the at least one data matrix, thereby identifying at least one of the datapoints as anomalous using a residuals matrix, the residuals matrix computed between the data matrix and an approximation of the data matrix, by applying a truncated singular value decomposition (SVD) to the data matrix.

34. The method of claim 32 or 33, wherein the collected data comprises image data.

35. The method of claim 32 or 33, wherein the collected data comprises engineering data, the method being applied to detect or predict a machine fault based on the at least one anomalous datapoint.

36. The method of claim 32 or 33, wherein the collected data comprises medical data.

37. A computer system comprising one or more computers programmed or otherwise configured to carry out the method steps of any preceding claim.

38. A computer program comprising program instructions for programming a computer or a set of computers to carry out the method steps of any of claims 1 to 36.

39. The computer program of claim 38, embodied in non-transitory computer-readable storage media.