

We claim :

1. A method for detecting compromised scopes in an online service, comprising:
 - receiving detection results of behaviors occurring on devices within the online service;
 - scoring each of the detection results based on how anomalous the respective behaviors are within the online service, wherein each of the detection results are scored based on a ratio between a subpopulation count and a population count for the respective behaviors within the online service;
 - excluding one or more of the scored detection results having one or more of a subpopulation count and a population count for the behaviors below one or more of a subpopulation threshold and a population threshold, respectively
 - organizing a remaining one or more of the scored detection results according to scopes;
 - applying multi-signal detection logic to a given scope from the scopes to produce a confidence score indicating whether the given scope is compromised;
 - determining whether to present an alert of the given scope being compromised based on comparing the confidence score to an alert threshold; and
 - in response to determining that the alert is to be presented of the given scope being compromised, generating and transmitting the alert.
2. The method of claim 1, wherein the given scope is associated with a given device or a given user account within the online service.
3. The method of claim 1, wherein scoring each of the detection results further comprises:
 - computing an identifier for a given detection result;
 - retrieving an aggregate count for the given detection result;
 - incrementing aggregate population counts for the given detection result;
 - calculating an anomaly score for the given detection result based on the aggregate population counts; and
 - associating the anomaly score with the given detection result to produce a scored detection result.
4. The method of claim 3, wherein the aggregate population counts include:
 - a raw population count, incremented for each observed detection result; and
 - at least one subpopulation count, incremented for each observed detection result that includes a characteristic defining a subpopulation of detection results.

5. The method of claim 3, wherein the identifier includes at least one field from the given detection result, including :
 - a role for a given device associated with the given detection result;
 - a user associated with the given detection result;
 - a behavior class associated with the given detection results;
 - and
 - a timestamp.
6. The method of claim 1, wherein organizing the remaining one or more of the scored detection results according to the scopes further comprises:
 - receiving a scope definition identifying a characteristic by which to divide the remaining one or more of the scored detection results;
 - receiving a time window from which the remaining one or more of the scored detection results are to be analyzed; and
 - dividing the remaining one or more of the scored detection results that were observed within the time window into the scopes according to values of the identified characteristic.
7. The method of claim 6, wherein the characteristic specifies values for :
 - device identifiers;
 - user accounts; and
 - process identifiers.
8. The method of claim 1, wherein detections results from a smaller window of time, comprising several minutes of detections results, and a larger window of time, comprising several days of detection results, are included in the scopes.
9. The method of claim 1 , wherein applying the multi-signal detection logic to the given scope to produce the confidence score further comprises:
 - selecting a predictive model;
 - extracting characteristics from a group of the remaining one or more of the scored detection results that were organized according to the given scope;
 - scoring the characteristics for conversion into numerically valued features; and
 - providing the features to the predictive model to generate the confidence score.
10. The method of claim 9, wherein the predictive model is generated and selected based on a continuous machine learning process.
11. A system including a processor and a memory storage device storing instructions that when executed by the processor provide for detecting compromised scopes in an online service, comprising:

a signature anomaly scorer, configured to:

- receive a given detection result of an event observed occurring in the online service;

- determine an extent to which a behavior associated with the given detection result has been previously observed as occurring in the online service by determining a subpopulation count and a population count for the behaviour in the online service;

- produce an anomaly score to be associated with the given detection result based on a ratio between the subpopulation count and the population count for the behavior; and

- determine the subpopulation count and the population count for the behaviour is above a subpopulation threshold and a population threshold respectively;

a multi-signal detector in communication with the signature anomaly scorer, configured to:

- receive one or more detection results sharing a scope in the online service with the given detection result;

- receive, from the signature anomaly scorer, the anomaly score associated with the given detection result;

- extract features from the given detection result and the one or more detection results sharing the scope, the features including the anomaly score; and

- generate a confidence score based on the extracted features for whether the scope is compromised.

12. The system of claim 11, wherein the one or more detection results sharing the scope in the online service with the given detection result are observed within a window of time.

13. The system of claim 12, wherein the extent to which the behavior associated with the given detection result has been previously observed as occurring in the online service is based on occurrences during the window of time.

14. The system of claim 11, wherein the extent to which the behavior associated with the given detection result has been previously observed as occurring in the online service is based on:

- a raw population count of events occurring in the online service; and
- a raw subpopulation count of a number of occurrences of the behavior in the online service.

15. The system of claim 11, wherein the extent to which the behavior associated with the given detection result has been previously observed as occurring in the online service is based on:

a scoped population count of a number of entities comprising a scope of the online service; and

a scoped subpopulation count of a number of entities in the scope associated with the behavior.

16. The system of claim 15, wherein the entities comprising the scope is defined by an analyst as one of devices within the online service or user accounts of the online service.

17. A computer readable storage device including processor executable instructions for detecting compromised scopes in an online service, comprising:

receiving detection results of behaviors occurring on devices within the online service;

scoring each of the detection results based on how anomalous the respective behaviors are within the online service, wherein each of the detection results are scored based on a ratio between a subpopulation count and a population count for the respective behaviors within the online service;

excluding one or more of the scored detection results having one or more of a subpopulation count and a population count for the respective behaviors below one or more of a subpopulation threshold and a population threshold, respectively;

organizing a remaining one or more of the scored detection results according to scopes, wherein a given scope from the scopes is associated with a given device or a given user account within the online service;

applying multi-signal detection logic to the given scope to produce a confidence score indicating whether the given scope is compromised;

determining whether to alert an analyst of the given scope being compromised based on comparing the confidence score to an alert threshold; and

in response to determining that the analyst is to be alerted of the given scope being compromised, generating and transmitting an alert.

18. The computer readable storage device of claim 17, wherein scoring each of the detection results further comprises :

computing an identifier for a given detection result;
retrieving an aggregate count for the given detection result;
incrementing aggregate population counts for the given detection result;
calculating an anomaly score for the given detection result based on the aggregate population counts; and

associating the anomaly score with the given detection result to produce a scored detection result.

19. The computer readable storage device of claim 17, wherein organizing the remaining one or more of the scored detection results according to the scopes further comprises :

receiving a scope definition identifying a characteristic by which to divide the remaining one or more of the scored detection results;

receiving a time window from which the remaining one or more of the scored detection results are to be analyzed;

dividing the remaining one or more of the scored detection results that were observed within the time window into the scopes according to values of the identified characteristic.

20. The computer readable storage device of claim 17, wherein applying the multi-signal detection logic to the given scope to produce the confidence score further comprises :

selecting a predictive model;

extracting characteristics from a group of the remaining one or more of the scored detection results that were organized according to the given scope;

scoring the characteristics for conversion into numerically valued features;

and

providing the features to the predictive model to generate the confidence score.