

CLAIMS

1. A server computer system comprising at least one hardware processor configured to:
 - in response to receiving an indication of an occurrence of a target event on a target client system, assemble an event sequence including the target event, all events of the event sequence having occurred on the target client system, wherein members of the event sequence are arranged according to a time of occurrence of each event of the event sequence;
 - in response to receiving the indication, select a parameter value from a plurality of parameter values according to the target client system;
 - in response selecting the parameter value, instantiate a behavior model with the parameter value, the behavior model configured to input a selected event of the event sequence and in response, to produce a prediction indicator indicative of a likelihood that the event sequence includes the target event; and
 - in response to instantiating the behavior model, determine whether the target event is indicative of a computer security threat according to the prediction indicator;wherein:
 - the plurality of parameter values are determined according to a training corpus of events having occurred on a plurality of client systems grouped into a plurality of client cluster;
 - each of the plurality of parameter values is determined according to a distinct subset of the training corpus, each distinct subset having occurred on members of a distinct cluster of a plurality of client clusters; and
 - selecting the parameter value comprises selecting a target cluster from the plurality of client clusters according to the target client system, and selecting the parameter value according to the target cluster.
2. The computer system of claim 1, wherein determining whether the target event is indicative of the computer security threat comprises:

3 comparing the prediction indicator to a pre-determined threshold;
4 in response, when the comparison indicates that the event sequence is likely to
5 include the target event, determining that the target event is not indicative
6 of the computer security threat; and
7 when the comparison indicates that the event sequence is not likely to include the
8 target event, determining that the target event is indicative of the computer
9 security threat.

10
1 3. The computer system of claim 1, wherein determining whether the target event is
2 indicative of the computer security threat comprises:
3 determining a predicted event according to the prediction indicator;
4 comparing the predicted event to the target event; and
5 in response, determining whether the target event is indicative of the computer
6 security threat according to a result of the comparison.

7
1 4. The computer system of claim 1, wherein:
2 the event sequence comprises a central event and an event context, the event
3 context comprising a first subset of the events occurring prior to the
4 central event and a second subset of events occurring later than the central
5 event;
6 the selected event is a member of the event context; and
7 the prediction indicator is indicative of a likelihood that the target event is the
8 central event of the event sequence.

9
1 5. The computer system of claim 1, wherein:
2 the behavior model comprises an event encoder configured to determine a set of
3 coordinates indicative of a position of the target event in an event
4 embedding space; and

5 the hardware processor is further configured to determine whether the target event
6 is indicative of the computer security threat according to the position.

7
1 6. The computer system of claim 1, wherein grouping the plurality of client systems into
2 clusters comprises:

3 constructing a plurality of event categories according to whether a training event
4 of the training corpus occurs in a selected context of other events of the
5 training corpus;

6 in response to constructing the plurality of event categories, deciding a cluster
7 assignment of each client systems according to an event profile of the each
8 client system, the event profile determined according to a count of events
9 occurring on the each client system and belonging to a selected event
10 category of the plurality of event categories.

11
1 7. The computer system of claim 1, wherein the parameter value is determined
2 according to a result of training another instance of the behavior model on events
3 having occurred on members of the target cluster, and wherein training the
4 another instance of the behavior model comprises:

5 assembling a training sequence of events having a central event and an event
6 context;

7 employing the another instance of the behavior model, in response to receiving
8 the event context of the training sequence, to produce a training prediction
9 indicative of a likelihood that the central event belongs to the training
10 sequence; and

11 determining the parameter value according to the training prediction.

12
1 8. The computer system of claim 1, further configured to select the target cluster
2 according to whether the target client system is a member of the target cluster.
3

1 9. The computer system of claim 1, wherein selecting the target cluster comprises:
2 determining a set of coordinates indicating a position of the target client system in
3 a multidimensional embedding space; and
4 in response, selecting the target cluster according to the set of coordinates.
5

1 10. The computer system of claim 9, configured to select the target cluster further
2 according to another set of coordinates indicating a position of a member
3 of the target cluster in the embedding space.
4

1 11. The computer system of claim 1, wherein the target event comprises a launch of a
2 selected process on the target client system.
3

1 12. A computer-implemented method comprising:
2 in response to receiving an indication of an occurrence of a target event on a target client
3 system, employing at least one hardware processor of a computer system to
4 assemble an event sequence including the target event, all events of the event
5 sequence having occurred on the target client system, wherein members of the
6 event sequence are arranged according to a time of occurrence of each event of
7 the event sequence;
8 in response to receiving the indication, employing at least one processor of the computer
9 system to select a parameter value from a plurality of parameter values according
10 to the target client system;
11 in response to selecting the set parameter value, employing at least one hardware
12 processor of the computer system to instantiate a behavior model with the
13 parameter value, the behavior model configured to input a selected event of the
14 event sequence and in response, to produce a prediction indicator indicative of a
15 likelihood that the event sequence includes the target event; and

16 in response to instantiating the behavior model, employing at least one hardware
17 processor of the computer system to determine whether the target event is
18 indicative of a computer security threat according to the prediction indicator;

19 wherein:

20 the plurality of parameter values are determined according to a training corpus of
21 events having occurred on a plurality of client systems grouped into a
22 plurality of client clusters;

23 each of the plurality of parameter values is determined according to a distinct
24 subset of the training corpus, each distinct subset having occurred on
25 members of a distinct cluster of a plurality of client clusters; and

26 selecting the parameter value comprises selecting a target cluster from the
27 plurality of client clusters according to the target client system, and
28 selecting the parameter value according to the target cluster.

1 13. The method of claim 12, wherein determining whether the target event is indicative of
2 the computer security threat comprises:

3 comparing the prediction indicator to a pre-determined threshold;

4 in response, when the comparison indicates that the event sequence is likely to
5 include the target event, determining that the target event is not indicative
6 of the computer security threat; and

7 when the comparison indicates that the event sequence is not likely to include the
8 target event, determining that the target event is indicative of the computer
9 security threat.

10 14. The method of claim 12, wherein determining whether the target event is indicative of
1 the computer security threat comprises:

2 determining a predicted event according to the prediction indicator;

3 comparing the predicted event to the target event; and
4

5 in response, determining whether the target event is indicative of the computer
6 security threat according to a result of the comparison.

7
1 15. The method of claim 12, wherein:

2 the event sequence comprises a central event and an event context, the event
3 context comprising a first subset of the events occurring prior to the
4 central event and a second subset of events occurring later than the central
5 event;

6 the selected event is a member of the event context; and

7 the prediction indicator is indicative of a likelihood that the target event is the
8 central event of the event sequence.

9
1 16. The method of claim 12, wherein:

2 the behavior model comprises an event encoder configured to determine a set of
3 coordinates indicative of a position of the target event in an event
4 embedding space; and

5 the at least one hardware processor is further configured to determine whether the
6 target event is indicative of the computer security threat according to the
7 position.

8
1 17. The method of claim 12, wherein grouping the plurality of client systems into clusters
2 comprises:

3 constructing a plurality of event categories according to whether a training event
4 of the training corpus occurs in a selected context of other events of the
5 training corpus;

6 in response to constructing the plurality of event categories, deciding a cluster
7 assignment of each client systems according to an event profile of the each
8 client system, the event profile determined according to a count of events

9 occurring on the each client system and belonging to a selected event
10 category of the plurality of event categories.
11

1 18. The method of claim 12, wherein the parameter value is determined according to a
2 result of training another instance of the behavior model on events having
3 occurred on members of the target cluster, and wherein training the another
4 instance of the behavior model comprises:
5 assembling a training sequence of events having a central event and an event
6 context;
7 employing the another instance of the behavior model, in response to receiving
8 the event context of the training sequence, to produce a training prediction
9 indicative of a likelihood that the central event belongs to the training
10 sequence; and
11 determining the parameter value according to the training prediction.
12

1 19. The method of claim 12, further comprising selecting the target cluster according to
2 whether the target client system is a member of the target cluster.
3

1 20. The method of claim 12, wherein selecting the target cluster comprises:
2 determining a set of coordinates indicating a position of the target client system in
3 a multidimensional embedding space; and
4 in response, selecting the target cluster according to the set of coordinates.
5

1 21. The method of claim 20, comprising selecting the target cluster further
2 according to another set of coordinates indicating a position of a member
3 of the target cluster in the embedding space.
4

1 22. The method of claim 12, wherein the target event comprises a launch of a selected
2 process on the target client system.

- 3
- 1 23. A non-transitory computer-readable medium storing instructions which, when executed
- 2 by at least one hardware processor of a computer system, cause the computer system to:
- 3 in response to receiving an indication of an occurrence of a target event on a target client
- 4 system, assemble an event sequence including the target event, all events of the
- 5 event sequence having occurred on the target client system, wherein members of
- 6 the event sequence are arranged according to a time of occurrence of each event
- 7 of the event sequence;
- 8 in response to receiving the indication, select a parameter value from a plurality of
- 9 parameter values according to the target client system;
- 10 in response selecting the parameter value, instantiate a behavior model with the
- 11 parameter value, the behavior model configured to input a selected event of the
- 12 event sequence and in response, to produce a prediction indicator indicative of a
- 13 likelihood that the event sequence includes the target event; and
- 14 in response to instantiating the behavior model, determine whether the target event is
- 15 indicative of a computer security threat according to the prediction indicator;
- 16 wherein:
- 17 the plurality of parameter values are determined according to a training corpus of
- 18 events having occurred on a plurality of client systems grouped into a
- 19 plurality of client clusters;
- 20 each of the plurality of parameter values is determined according to a distinct
- 21 subset of the training corpus, each distinct subset having occurred on
- 22 members of a distinct cluster of a plurality of client clusters; and
- 23 selecting the parameter value comprises selecting a target cluster from the
- 24 plurality of client clusters according to the target client system, and
- 25 selecting the parameter value according to the target cluster.
- 26

CLAIMS

1. A server computer system comprising at least one hardware processor configured to:
 - in response to an occurrence of a target event on a target client system, the target event selected from a group consisting of a launch of a process and an attempt to access a file, determine according to the target event whether the target client system comprises malicious software; and
 - in response to a determination that the target client system comprises malicious software, perform a security action to protect the target client system;
 - wherein determining whether the target client system comprises malicious software comprises:
 - determining an event context of the target event, the event context comprising a set of events having occurred on the target client system prior to the target event and another set of events having occurred on the client system following the target event,
 - selecting a client profile from a plurality of pre-determined client profiles according to the target event, the selected client profile characterizing a baseline behavior of a selected cluster of client systems,
 - in response to selecting the client profile, employing a behavior model specific to the selected client profile to determine a plurality of prediction scores according to the event context, each score of the plurality of prediction scores determined for a distinct event type and indicative of a likelihood that an event of the distinct event type would occur in the event context for the selected client profile,
 - selecting, from the plurality of prediction scores, a score determined for an event type of the target event, and
 - comparing the selected score to a pre-determined threshold; and
 - wherein a composition of the selected cluster of client systems is determined by analyzing a collection of events having occurred on a plurality of client systems to determine which of the plurality of client systems show similar behavior.

2. The computer system of claim 1, wherein determining whether the target client system comprises malicious software comprises:

when the selected score indicates that the target event is likely to occur in the event context, determining that the target client system does not comprise malicious software; and

when the selected score indicates that the target event is not likely to occur in the event context, determining that the target client system comprises malicious software.

3. The computer system of claim 1, wherein determining whether the target client system comprises malicious software comprises:

determining a predicted event according to the plurality of prediction scores;

comparing the predicted event to the target event; and

in response, determining whether the target client system comprises malicious software according to a result of the comparison.

4. The computer system of claim 1, wherein:

the behavior model comprises an event encoder configured to determine a set of coordinates indicative of a position of the target event in an event embedding space; and

the hardware processor is further configured to determine whether the target client system comprises malicious software according to the position.

5. The computer system of claim 1, wherein determining the composition of the selected cluster of client systems comprises:

constructing a plurality of event categories according to whether a selected event of the collection of events occurs in a selected context of other events of the collection of events;

in response to constructing the plurality of event categories, deciding whether each client system belongs to the selected cluster of client systems according to an event profile of the each client system, the event profile determined according to a count of events occurring on the each client system and belonging to a selected event category of the plurality of event

categories.

6. The computer system of claim 1, wherein training the behavior model comprises:

assembling a training sequence of events having a training central event and a training event context, wherein each member of the training sequence is selected from the collection of events according to whether the each member has occurred on a member of the selected cluster of client systems;

employing the behavior model, in response to receiving the training event context of the training sequence, to produce a training prediction indicative of a likelihood that the training central event belongs to the training sequence; and

adjust a set of parameter values of the behavior model according to the training prediction.

7. The computer system of claim 1, further configured to select the client profile according to whether the target client system is a member of the selected cluster.

8. The computer system of claim 1, wherein selecting the client profile comprises:

determining a set of coordinates indicating a position of the target client system in a multidimensional embedding space;

in response, selecting a target cluster of the plurality of clusters according to the set of coordinates; and

selecting the client profile according to the target cluster.

9. The computer system of claim 8, configured to select the target cluster further according to another set of coordinates indicating a position of a member of the target cluster in the embedding space.

10. A computer-implemented method comprising:

in response to an occurrence of a target event on a target client system, the target event selected from a group consisting of a launch of a process and an attempt to access a file,

employing at least one hardware processor of a computer system in determining according to the target event, whether the target client system comprises malicious software; and

in response to a determination that the target client system comprises malicious software, performing a security action to protect the target client system;

wherein determining whether the target client system comprises malicious software comprises:

determining an event context of the target event, the event context comprising a set of events having occurred on the target client system prior to the target event and another set of events having occurred on the client system following the target event,

selecting a client profile from a plurality of pre-determined client profiles according to the target event, the selected client profile characterizing a baseline behavior of a selected cluster of client systems,

in response to selecting the client profile, employing a behavior model specific to the selected client profile to determine a plurality of prediction scores according to the event context, each score of the plurality of prediction scores determined for a distinct event type and indicative of a likelihood that an event of the distinct event type would occur in the event context for the selected client profile,

selecting, from the plurality of prediction scores, a score determined for an event type of the target event, and

comparing the selected score to a pre-determined threshold; and

wherein a composition of the selected cluster of client systems is determined by analyzing a collection of events having occurred on a plurality of client systems to determine which of the plurality of client systems show similar behavior.

11. The method of claim 10, wherein determining whether the target client system comprises malicious software comprises:

when the selected score indicates that the target event is likely to occur in the event context, determining that the target client system does not comprise malicious software; and

when the selected score indicates that the target event is not likely to occur in the event context, determining that the target client system comprises malicious software.

12. The method of claim 10, wherein determining whether the target client system comprises malicious software comprises:

determining a predicted event according to the plurality of prediction scores;

comparing the predicted event to the target event; and

in response, determining whether the target client system comprises malicious software according to a result of the comparison.

13. The method of claim 10, wherein:

the behavior model comprises an event encoder configured to determine a set of coordinates indicative of a position of the target event in an event embedding space; and

the at least one hardware processor is further configured to determine whether the target client system comprises malicious software according to the position.

14. The method of claim 10, wherein determining the composition of the selected cluster of client systems comprises:

constructing a plurality of event categories according to whether a selected event of the collection of events occurs in a selected context of other events of the collection of events;

in response to constructing the plurality of event categories, deciding whether each client system belongs to the selected cluster of client systems according to an event profile of the each client system, the event profile determined according to a count of events occurring on the each client system and belonging to a selected event category of the plurality of event categories.

15. The method of claim 10, wherein training the behavior model comprises:

assembling a training sequence of events having a training central event and a training event context, wherein each member of the training sequence is selected from the collection

of events according to whether the each member has occurred on a member of the selected cluster of client systems;

employing the behavior model, in response to receiving the training event context of the training sequence, to produce a training prediction indicative of a likelihood that the training central event belongs to the training sequence; and

adjusting a set of parameters of the behavior model according to the training prediction.

16. The method of claim 10, further comprising selecting the client profile according to whether the target client system is a member of the selected cluster.

17. The method of claim 10, wherein selecting the client profile comprises:

determining a set of coordinates indicating a position of the target client system in a multidimensional embedding space;

in response, selecting a target cluster of the plurality of clusters according to the set of coordinates; and

selecting the client profile according to the target cluster.

18. The method of claim 17, comprising selecting the target cluster further according to another set of coordinates indicating a position of a member of the target cluster in the embedding space.

19. A non-transitory computer-readable medium storing instructions which, when executed by at least one hardware processor of a computer system, cause the computer system to:

in response to an occurrence of a target event on a target client system, the target event selected from a group consisting of a launch of a process and an attempt to access a file, determine according to the target event whether the target client system comprises malicious software; and

in response to a determination that the target client system comprises malicious software, perform a security action to protect the target client system;

wherein determining whether the target client system comprises malicious software comprises:

determining an event context of the target event, the event context comprising a set of events having occurred on the target client system prior to the target event and another set of events having occurred on the client system following the target event,

selecting a client profile from a plurality of pre-determined client profiles according to the target event, the selected client profile characterizing a baseline behavior of a selected cluster of client systems,

in response to selecting the client profile, employing a behavior model specific to the client profile to determine a plurality of prediction scores according to the event context, each score of the plurality of prediction scores determined for a distinct event type and indicative of a likelihood that an event of the distinct event type would occur in the event context for the selected client profile,

selecting, from the plurality of prediction scores, a score determined for an event type of the target event, and

comparing the selected score to a predetermined threshold; and

wherein a composition of the selected cluster of client systems is determined by analyzing a collection of events having occurred on a plurality of client systems to determine which of the plurality of client systems show similar behavior.

20. The computer system of claim 1, wherein the at least one hardware processor is configured to select the client profile further according to a current user of the target client system.

21. The method of claim 10, wherein the at least one hardware processor is configured to select the client profile further according to a current user of the target client system.

22. The computer system of claim 1, wherein the event context of the target event

comprises an attempt by the target client system to access a selected network location.

23. The method of claim 10, wherein the event context of the target event comprises an attempt by the target client system to access a selected network location.