

CLAIMS

1. A server computer system comprising at least one hardware processor configured to execute an event encoder, an event clustering engine and a client clustering engine, wherein:

the event encoder is trained to determine a position of each event of a training corpus in an event embedding space, the training corpus comprising a collection of events having occurred on a plurality of client systems;

the event clustering engine is configured to distribute the training corpus among a plurality of event clusters according to the position of each event in the event embedding space; and

the client clustering engine is configured to:

distribute the plurality of client systems among a plurality of client clusters according to a content of each event cluster of the plurality of event clusters, and

in response to distributing client systems to client clusters, transmit a client cluster membership indicator to an anomaly detector configured to determine whether a target event occurring on a target client system is indicative of a computer security threat according to a subset of the training corpus consisting of events having occurred on members of a target client cluster selected from the plurality of client clusters to include the target client system;

wherein the event encoder is configured to determine a position of a selected event according to an event sequence comprising the selected event, a first event, and a second event of the training corpus, the first event having occurred earlier than the selected event, and the second event having occurred later than the selected event; and

wherein training the event encoder comprises:

coupling the event encoder to an event decoder configured to receive an indicator of the position of the selected event and to output a prediction indicator indicative of a likelihood that the event sequence comprises the first and second events, and

adjusting a set of parameters of the event encoder according to the prediction indicator.

2. The computer system of claim 1, wherein training the event encoder further comprises:

employing the event decoder to output a first predicted event and a second predicted event according to the position of the selected event;

comparing the first predicted event to the first event and the second predicted event to the second event; and

adjusting the set of parameters of the event encoder according to a result of the comparison.

3. The computer system of claim 1, wherein the client clustering engine is configured to assign a selected client system to a selected client cluster according to an event profile of the selected client system, the event profile comprising a plurality of components, each distinct component determined according to a count of events having occurred on the selected client system and belonging to a distinct event cluster of the plurality of event clusters.

4. The computer system of claim 3, wherein distributing the plurality of client systems among client clusters comprises assigning client systems having similar event profiles to the same client cluster.

5. The computer system of claim 1, wherein an event of the training corpus comprises a launch of a selected process on a client system of the plurality of client systems.

6. A computer-implemented method comprising employing at least one hardware processor of a computer system to:

execute an event encoder trained to determine a position of each event in an event embedding space, the training corpus comprising a collection of events having occurred on a plurality of client systems;

distribute the training corpus among a plurality of event clusters according to the position of each event in the event embedding space;

in response to distributing events to event clusters, distribute the plurality of client systems among a plurality of client clusters according to a content of each event cluster of the plurality of event clusters; and

in response to distributing client systems to client clusters, transmit a client cluster membership indicator to an anomaly detector configured to determine whether a target event occurring on a target client system is indicative of a computer security threat according to a subset of the training corpus consisting of events having occurred on members of a target client cluster selected from the plurality of client clusters to include the target client system;

wherein the event encoder is configured to determine a position of a selected event according to an event sequence comprising the selected event, a first event, and a second event of the training corpus, the first event having occurred earlier than the selected event, and the second event having occurred later than the selected event; and

wherein training the event encoder comprises:

coupling the event encoder to an event decoder configured to receive an indicator of the position of the selected event and to output a prediction indicator indicative of a likelihood that the event sequence comprises the first and second events, and

adjusting a set of parameters of the event encoder according to the prediction indicator.

7. The method of claim 6, wherein training the event encoder further comprises:

employing the event decoder to output a first predicted event and a second predicted event according to the position of the selected event;

comparing the first predicted event to the first event and the second predicted event to the second event; and

adjusting the set of parameters of the event encoder according to a result of the comparison.

8. The method of claim 6, comprising assigning a selected client system to the selected client cluster according to an event profile of the client system, the event profile comprising a plurality of components, each distinct component determined according to a count of events having occurred on the selected client system and belonging to a distinct event cluster of the plurality of event clusters.

9. The method of claim 8, wherein distributing the plurality of client systems among client clusters comprises assigning client systems having similar event profiles to the same client cluster.

10. The method of claim 6, wherein an event of the training corpus comprises a launch of a selected process on a client system of the plurality of client systems.

11. 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 execute an event encoder, an event clustering engine and a client clustering engine, wherein:

the event encoder is trained to determine a position of each event of a training corpus in an event embedding space, the training corpus comprising a collection of events having occurred on a plurality of client systems;

the event clustering engine is configured to distribute the training corpus among a plurality of event clusters according to the position of each event in the event embedding space; and

the client clustering engine is configured to:

distribute the plurality of client systems among a plurality of client clusters according to a content of each event cluster of the plurality of event clusters; and

in response to distributing client systems to client clusters, transmit a client cluster membership indicator to an anomaly detector configured to determine whether a target event occurring on a target client system is indicative of a computer security threat according to a subset of the training corpus consisting of events having occurred on members of a target client cluster selected from the plurality of client clusters to include the target client system;

wherein the event encoder is configured to determine a position of a selected event according to an event sequence comprising the selected event, a first event, and a second event of the training corpus, the first event having occurred earlier than the selected event, and the second event having occurred later than the selected event; and

wherein training the event encoder comprises:

coupling the event encoder to an event decoder configured to receive an indicator of the position of the selected event and to output a prediction indicator indicative of a likelihood that the event sequence comprises the first and second events, and

adjusting a set of parameters of the event encoder according to the prediction indicator.