CLAIMS

1. A system (200) for processing a data package (100), the system comprising:

an input module (220) for receiving a data package (100) over a data network (230), the data package (100) comprising a header (110) and payload data (120);

storage (240) for storing at least one data model;

a buffer (260) for temporarily storing at least a part of the data package (100);

an analysis module (252) for analysing, at a first time, the header (110) of the data package (110) based on a plurality of characteristics, the characteristics comprising at least an identifier and a transmission time, and for determining, using the at least one data model, at least one probability value associated with a relationship between the plurality of characteristics;

a decision-making module (250) for determining a risk score based on the at least one the probability value and the at least a part of the data package (100), the risk score being indicative of a handling action to apply to the data package (100);

a monitoring unit (272) for determining whether to update the at least one data model at the first time based at least in part on whether the risk score is being determined at the first time; and

an updating module (270) for recalculating the at least one data model in the storage (240), wherein the updating module is configured to:

update the at least one data model based on the data package when it is determined that the at least one data model should be updated at the first time; and

retrieve the at least one data package from the temporary storage, and updating, at a second time later than the first time, the at least one data model, when it is determined that the at least one data model should not be updated at the first time, and the risk score has been generated at the first time; and

a processing module (290) for applying the handling action to the data package (100).

2. The system (200) for processing a data package (100) according to claim 1, wherein the characteristics of the data package (100) include any of a domain, a sender identifier, a recipient identifier, a transmission time, a location, an IP address, a classification, device information, client information, metadata and content of the payload data (120).

3. The system (200) for processing a data package (100) according to claim 1 or claim 2, wherein the decision-making module (250) comprises:

at least one attribute checker (252) associated with one of the at least one data model, the probability value output by each data model being associated with an attribute of the header (110), and

a combination unit (254) for determining the risk score based on the plurality of probability values generated by the at least one attribute checker.

4. The system (200) for processing a data package (100) according to claim 3, wherein the attribute checker (252) generates the plurality of probability values using at least one machine learning function.

5. The system (200) for processing a data package (100) according to any previous claim, wherein the storage (240) is any of non-volatile memory, remote storage, and distributed storage.

6. The system (200) for processing a data package (100) according to any previous claim, wherein the buffer (260) is any of random-access memory, dynamic random-access memory, and volatile memory, non-volatile memory, read only memory or a solid-state memory.

7. The system (200) for processing a data package (100) according to any previous claim, wherein the updating module (270) further updates the data model based on a relationship between the header (110) of the data package (100) and data retrieved from at least one external information repository (280).

8. The system (200) for processing a data package (100) according to claim 7, wherein the external information repository (280) comprises data from one or more additional data services stored remotely to the system (200).

9. The system (200) for processing a data package (100) according to any previous claim, wherein the processing module (290) comprises an encryption unit for encrypting the payload data (120) of the data package (100), wherein a level of encryption is dependent upon the handling action.

10. A method (300) for processing a data package (100), the method (300) comprising the steps of:

receiving (310) a data package, the data package (100) comprising a header (110) and payload data (120);

temporarily (330) storing at least part of the data package (100);

analysing, at a first time, the header (110) of the data package (100) based on a plurality of characteristics, the characteristics comprising at least an identifier and a transmission time, and determining, using at least one data model, at least one probability value associated with a relationship between the plurality of characteristics;

determining (320) a risk score based on at least one criterion and the at least a part of the data package (100), the risk score being indicative of a handling action to apply to the data package (100);

determining whether to update the at least one data model at the first time based at least in part on whether the risk score is being determined at the first time, and based on the determination:

updating the at least one data model based on the data package when it is determined that the at least one data model should be updated at the first time; and

retrieving the at least one data package from temporary storage, and updating, at a second time later than the first time, the at least one data model, when it is determined that the at least one data model should not be updated at the first time, and the risk score has been generated at the first time; and

applying the handling action to the data package (100).

11. The method (300) for processing a data package (100) according to claim 10, wherein the characteristics of the data package (100) include any of a domain, a sender identifier, a recipient identifier, a transmission time, a location, an IP address, a classification, device information, client information, metadata and content of the payload data (120).

12. The method (300) for processing a data package (100) according to claim 10 or claim 11, wherein the step of determining (320) the risk score comprises:

outputting a plurality of probability values associated with the data model, each probability value associated with an attribute of the header (110); and

determining a risk score based on the plurality of probability values, by combining the plurality of probability values.

13. The method (300) for processing a data package (100) according to claim 12, wherein the plurality of probability values is generated using at least one machine learning function.

14. The method (300) for processing a data package (100) according to any of claims 10 to 13, wherein updating the at least one data model comprises:

generating an updated data model based on a relationship between the header (110) of the data package (100) and data retrieved from at least one external information repository.

15. The method (300) for processing a data package (100) according to claim 14, wherein the external information repository comprises data from one or more additional data services.

16. The method (300) for processing a data package according to any of claims 10 to 15, wherein the step of applying the handling action comprises encrypting the payload data (120) of the data package (100) wherein a level of encryption is dependent upon the handling action.

17. A non-transitory computer readable storage medium comprising a set of computer-readable instructions stored thereon which, when executed by at least one processor cause the processor to process a data package (100), the instructions comprising:

receiving (310) a data package (100), the data package (100) comprising a header (110) and payload data (120);

temporarily storing (330) at least a part of the data package (100);

analysing, at a first time, the header (110) of the data package (100) based on a plurality of characteristics, the characteristics comprising at least an identifier and a transmission time, and determining, using at least one data model, at least one probability value associated with a relationship between the plurality of characteristics;

determining (320) a risk score based on at least one criterion and the at least a part of the data package (100), the risk score being indicative of a handling action to apply to the data package (100);

determining whether to update the at least one data model at the first time based at least in part on whether the risk score is being determined at the first time, and based on the determination:

updating the at least one data model based on the data package when it is determined that the at least one data model should be updated at the first time; and

retrieving the at least one data package from temporary storage, and updating, at a second time later than the first time, the at least one data model, when it is determined that the at least one data model should not be updated at the first time, and the risk score has been generated at the first time; and

applying the handling action to the data package (100).