Complexity, sophistication, and rate of growth of modern networks, coupled with the depth, continuity, and pervasiveness of their role in our everyday lives, stress the importance of identifying potential misuse or threats that could undermine regular operation.

To ensure an adequate and prompt reaction, anomalies in network traffic should be detected, classified, and identified as quickly and correctly as possible.

Several approaches focus on inspecting the content of packets traveling through the network, while other techniques aim at detecting suspicious activity by measuring the network state and comparing it with an expected baseline.

Formalizing a model for normal behavior requires the collection and analysis of traffic, in order to isolate a set of features capable of describing traffic completely and in a compact way.

The main focus of this dissertation is the quest for good representations for network traffic, representation that are abstract and can capture and describe much of the intricate structure of observed data in a simple manner.

In this way, some of the hidden factors and variables governing the traffic data generation process can be unveiled and disentangled and anomalous events can be spotted more reliably.

We adopted several methods to achieve such simpler representations, including Independent Component Analysis and deep learning architectures.

Machine learning techniques have been used for verifying the improvement in classification effectiveness that can be achieved with the proposed representations.