# An Ontology-based Recommendation System for Context-aware Network Monitoring

Ricardo F. Silva[1], Paulo Carvalho[1], Solange Rito Lima[1], Luis Álvarez Sabucedo[2], Juan M. Santos Gago[2], and João Marco C. Silva[3]

[1] Centro Algoritmi, Universidade do Minho, 4710-057 Braga, Portugal
{pmc,solange}@di.uminho.pt
[2] University of Vigo, Dept. of Telematics, Vigo, Spain
{Luis.Sabucedo,Juan.Gago}@det.uvigo.es
[3] HASLab, INESC TEC, Universidade do Minho, Braga, Portugal
joao.marco@inesctec.pt

**Abstract.** Current network management systems urge for a context-aware perspective of the provided network services and the underlying infrastructure usage. This need results from the heterogeneity of services and technologies in place, and from the massive traffic volumes traversing today's networks. To reduce complexity and improve interoperability, monitoring systems need to be flexible, context-aware, and able to self-configure measurement points (MPs) according to network monitoring tasks requirements. In addition, the use of sampling techniques in MPs to reduce the amount of traffic collected, analysed and stored has become mandatory and, currently, distinct sampling schemes are available for use in operational environments.
In this context, the main objective of this paper is the ontological definition of measurement requirements and components in sampling-based monitoring environments, with the aim of supporting an expert recommendation system able to understand context and identify the appropriate configuration rules to apply to a selection of MPs. In this way, the ontology, defining management needs, network measurement topology and sampling techniques, is described and explored considering several network management activities. A use case focusing on traffic accounting as monitoring task is also provided, demonstrating the expressiveness of the ontology and the role of the recommendation system in assisting context-aware network monitoring based on traffic sampling.

**Keywords:** Ontology; Network monitoring; Traffic sampling

## 1 Introduction

The semantic support for services, in a broad sense, is a common feature in many platforms nowadays. The features unleashed by semantic tools are achieving a high-maturity level, and their support for add-value services are the reasons for its broad adoption in a large number of scopes. Nevertheless, its adoption in traffic monitoring scenarios has yet a substantial way to evolve, especially, in heterogeneous and context-dependent environments, such as Smart Cities.

As a crucial task supporting network management activities, network monitoring must attend to each specific context requiring traffic measurements. This means that a system can use important context information to provide customised and optimised measuring services to meet the needs of network users/administrators. Furthermore, context-aware monitoring enables saving computational and communication resources, thus fostering the provision of more relevant and agile services. Due to the constant increase in data volumes, it is also essential to enhance the monitoring process without compromising its efficiency. For this purpose, the use of traffic sampling techniques is mandatory to enable capturing the behaviour of services and networks resorting uniquely to a subset of traffic [1].

In terms of context-representation requirements, several considerations must be taken into account when selecting the technologies that can satisfy a context middleware. The choice of how to store, represent and infer context in a context processor are important requirements for the operation of a context-based system. The use of ontological support allows to formalise semantically the interaction between the users and the machine, allowing reutilisation principles which facilitate the process of representing the information [2].

This paper handles the issue of defining an ontology to assist context-aware monitoring environments which resort to traffic sampling to improve monitoring efficiency. Using a highly practical approach, the present work contributes for the identified objective through: (i) the definition of a context-aware monitoring system architecture and associated recommendation system, capable of mapping measurement needs into a set of rules to configure measurement points (MPs); and (ii) the definition of the corresponding ontology, expressing relevant classes, relations and attributes of a monitoring task, the underlying network measurement topology and the available sampling techniques, which supports the configuration of the sampling-based monitoring environment. The semantic validation of this proposal is achieved through the application of the ontology in several management competence queries and, in particular, when supporting traffic accounting. For this purpose, the ontology has been previously populated with real data collected from the University of Minho campus network.

The first step to achieve the high-level design goals mentioned above is the provision of ontological support. Thus, after discussing related work in Section 2, a context-aware monitoring architecture is proposed in Section 3. The ontological layer, fully described in Section 4, must endow the maximum possible level of interoperability among components, enabling the deployment of autonomic network monitoring. This section formalises the domain concepts in terms of their relations and attributes in order to assist the deployment of added-value monitoring services in this context. A description of technologies involved in the process is also included. This semantic layer is intended to cope with requirements of sampling-based monitoring. The latter are presented under the form of competence questions in a case study, described in Section 5, as an initial validation of the semantic model of the system. Finally, the main conclusions are presented to the user in Section 6, where the reader can access insight information about the proposed model, its current and future features in the scope of semantic-based expert systems in the domain of network monitoring.

## 2 Related work

Mapping network measurement requirements into the most suitable MP, traffic sampling technique, and underlying operation parameters have been topics of research along the last years. Globally, such efforts have identified different strategies able to provide high accurate results in manifold tasks, such as traffic classification and characterisation [1, 3, 4], SLA compliance [5, 6], QoS monitoring [7, 8], and network security [9, 10]. However, as network traffic is heavily dynamic and heterogeneous, a sampling solution used to estimate a particular parameter correctly may not be adequate for a different parameter or traffic type [11]. This creates the need of having previous and detailed knowledge about the monitored traffic as well as direct intervention of network administrators for tuning the measurement process according to monitoring requirements.

High-level recommendation systems based on ontologies emerge as a new mechanism to improve the semantic expressiveness of management activities, being a valuable approach to overcome the challenges mentioned above. In this way, some works have been addressing interoperability issues, where ontologies are used to map network managed objects defined in information models, such as SMI, GMDO, MIF, and IPFIX [12, 13].

Considering network management activities, related research is exploring ontological representation as a mechanism for supporting autonomic networks, in particular, for automated configuration [14, 15]. More specific works which require traffic measurements are mainly focused on QoS monitoring [16, 17] and network security [18, 19].

Although being considered a key enabler within network semantic management, exploiting ontology's capabilities to face the challenges of selecting the most suitable sampling-based monitoring strategy in context-dependent network environments is still an open issue.

## 3 Context-aware monitoring architecture

The proposed context-aware monitoring architecture, represented in Figure 1, illustrates the expert recommendation system as a key component for assisting network management tasks. At upper level, from a management plane perspective, each service or network management task is expected to specify, and subsequently meet, particular measuring requirements. These requirements specificity will be handled within the control plane, where the expert recommendation system will act downstream to suggest adequate configuration parameters for a set of measurement points (MPs). At lower level, MPs will collect and report traffic descriptors as traffic traces or datasets. Thus, the recommendation system receives as main inputs the specification of measurement requirements and traffic data, and will include the required ontological support and reasoning components.

Depending on the monitoring context and on network traffic variability, the system is expected to suggest a configuration profile so that the monitoring task can be efficiently accomplished. Efficiency is here understood as a trade-off between measurements accuracy and overhead.

The main modules of the semantic recommendation system are illustrated in Figure 2. The module called ML incorporates machine learning techniques for identifying

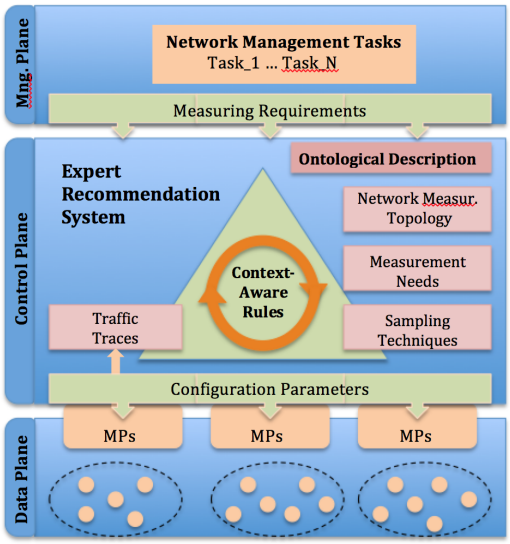


**Fig. 1.** Monitoring Architecture

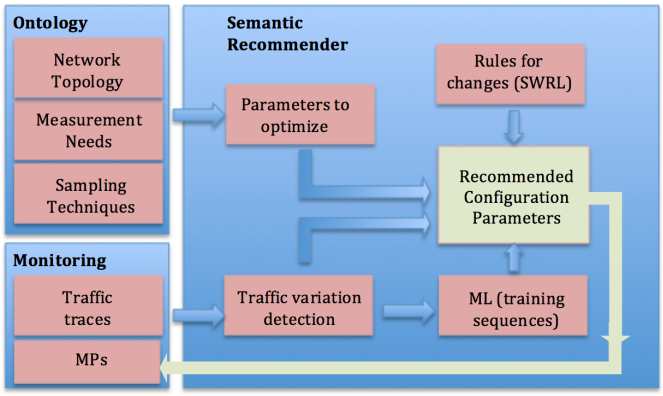


**Fig. 2.** Semantic Recommender

traffic behaviours from real data. The *Traffic variation detection* module detects traffic fluctuations and assesses when these may require changes on monitoring configuration. The module called *Rules for changes* is a repository of semantic rules (e.g., specified in SWRL) that indicate what changes must be accomplished in the process of monitoring a task. The module *Parameters to optimize* handles the parameters to be changed for a given monitoring configuration. As output, the system produces a specific configuration profile to be applied to a set of MPs.

## 4 Ontology definition

This section will explain the building steps of the ontology sustaining the monitoring architecture.

### 4.1 Conceptual model

Prior to the ontology definition, the preliminary step is the identification of the application domain and the specification of the questions of competence, i.e., the questions to which the ontological system is expected to answer. The purpose of the ontology is to serve as the basis for implementing a monitoring service. Therefore, the next step is to identify the concepts in the domain that should be represented in the ontology taking into account the specifications of the monitoring architecture.

### 4.2 Competence questions

Competence questions play a very important role both in the creation of an ontology since they allow to justify the existence of the ontology, and in the consequent evaluation of the ontology. When creating the ontology, the questions of competence must be verified so that the development of the ontology does not deviate from the purpose initially defined. Examples of competence questions are: (a) how much memory and CPU are available at a particular MP?; (b) which MPs are border routers?; (c) what are the characteristics of a particular MP?; (d) what is the id of the existing monitoring tasks? (e) which MPs are available in the network?; (f) what are the sampling techniques supported by a particular MP?; (g) what are the MPs capable of supporting a specific sampling technique?, or (h) what is the active technique and setup parameters at a particular MP?.

### 4.3 Ontology representation

Figure 3 illustrates the class model developed for the ontology, along with the corresponding object properties.

The monitoring architecture components *Management Needs*, *Measurement Topology* and *Sampling* are defined as classes in the ontology. This ontological representation allows a comprehensive understanding of the domain, establishing how these classes and the object properties are interrelated in a service monitoring context based on traffic

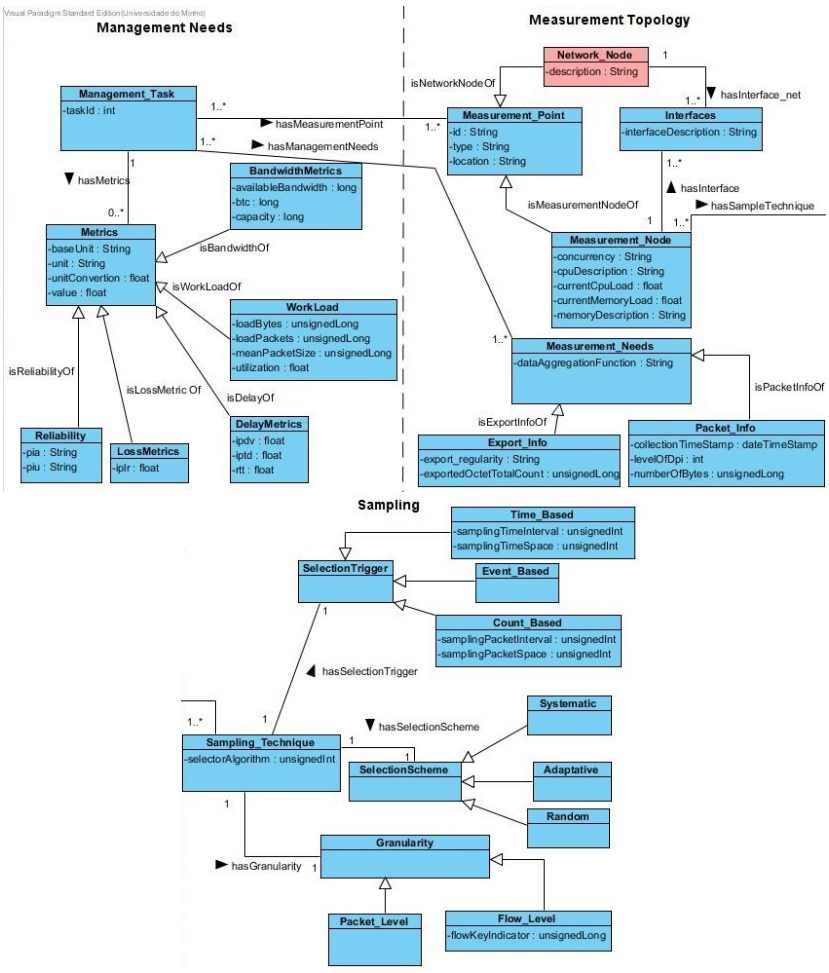


**Fig. 3.** Ontology: Model of classes

sampling. As illustrated, a `Management_Task`, such as accounting or traffic classification, determines the `Measurement_Needs`, in terms of `Metrics` to be measured, and a set of `Measurement_Point`. An MP can be a general purpose `Network_Node` (e.g., router or switch) performing network measurements or a dedicated `Measurement_Node`. Each `Measurement_Node` determines the sampling techniques supported in that node. These techniques are defined through the classes `SelectionTrigger`, `SelectionScheme` and `Granularity`, which determine the time and space characteristics regulating traffic sampling. In sampling techniques, timers, packet counters or events may trigger the sampling process, ruling the intervals in which packets are collected. Sampling intervals can be systematic, random or adaptive according to the current network load.

The model of classes also includes the attributes (data properties) of each object, e.g. `loadBytes`, `loadPackets`. Defining a management task, e.g. accounting, corresponds to an object instantiation (individual), which has object properties such as `hasMeasurementPoint` and `taskID`.

## 5 Case studies - Semantic validation

In this section, as a semantic validation of the devised ontology expressiveness, examples of real uses of the ontology in the context of network management are presented, as well as a study on the requirements of a specific management task, i.e., *traffic accounting*. The data used to populate the ontology was collected from the University of Minho network during a working day.

### 5.1 Querying the ontology

As mentioned, the ontology must be able to answer a list of competence questions. These questions ground the existence of the ontology and allow to evaluate whether the ontology responds to the defined purposes. To achieve this, the SPARQL Protocol and RDF Query Language (SPARQL), based on the Resource Description Framework (RDF), is used.

As an example, we only include a competence question from Section 4.2, and the corresponding answer. Other queries are covered within traffic accounting task.

- How much memory and CPU are available at a particular MP?
  *Input of Query 1:*

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX sm: <http://www.semanticweb.org/kaiser/ontologies/2016/3/
service-monitoring#>
SELECT ?nome ?memoria ?cpu
WHERE{
?mp rdf:type sm:Measurement_Point.
?mn rdf:type sm:Measurement_Node.
?mp sm:id ?nome.
?mp sm:hasMeasurementNode ?mn.
?mn sm:currentMemoryLoad ?memoria.
?mn sm:currentCpuLoad ?cpu
}
ORDER BY ?nome
```

| name | memory | cpu |
|---|---|---|
| "Point1"@ | "76566"^^<http://www.w | "5.03"^^<http://www.w3 |
| "Point2"@ | "96410"^^<http://www.w | "17.95"^^<http://www.w |
| "Point3"@ | "86163"^^<http://www.w | "18.26"^^<http://www.w |
| "Point4"@ | "80765"^^<http://www.w | "10.76"^^<http://www.w |
| "Point5"@ | "85551"^^<http://www.w | "97.27"^^<http://www.w |

**Fig. 4.** Available memory and CPU at MPs

In the example of Query 1, five variables are created: `mp` of type `Measurement_Point` (line 9); `mn` of type `Measurement_Node` (line 10); `name` that collects the values of the `id` attribute associated with each `Measurement_Point` (line 11); `memory` that collect the values of the `currentMemoryLoad` attribute associated with each `Measurement_Node` (line 13); and `cpu` that collects the values of the `currentCpuLoad` attribute associated with each `Measurement_Node` (line 14). The variables selected for the query output are `name`, `memory` and `cpu`, and corresponding values are shown in Figure 4.

### 5.2 Traffic accounting

Traffic accounting, a vital network management task for service providers, was selected as case study for testing the ontological system. In its simplest form, this task keeps track of users traffic volumes, including the amount of time spent per session, the amount of data transferred during a session, and the type of services accessed.

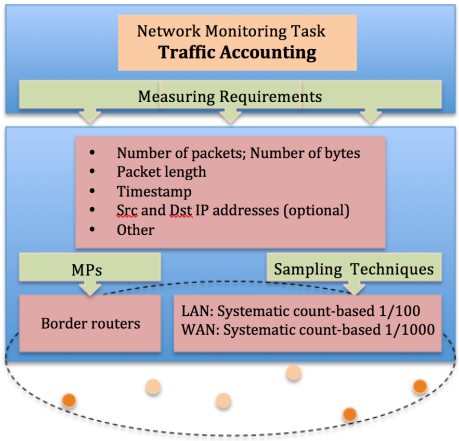


**Fig. 5.** Accounting task specification

Figure 5 illustrates essential aspects to consider when performing traffic accounting, namely: the measuring requirements; the MPs involved in traffic accounting; and the most adequate sampling techniques to use, i.e., the techniques that achieve a better

trade-off between accuracy and overhead (allowing to obtain accurate results without interfering with normal network operation). As shown, common requirements of traffic accounting include: the amount of traffic monitored (in packets and bytes); the size of the collected packets; a time instant when a sample is collected; the source and destination IP addresses; and other specific header (or payload) fields under observation. The selection of the most suitable MPs for the execution of accounting are, usually, the border routers of the network domain, but can also be dedicated measurement nodes, as defined in Section 4.

The selection of the sampling technique that best suits a monitoring task depends on several factors, such as the memory and CPU available at the MP, the amount of data collected, the network performance, etc. Sampling techniques with best experimental results in performing accounting are: systematic count-based 1/100 for local area networks, and systematic count-based 1/1000 for non-local networks [20].

Next, examples of queries and results obtained when performing traffic accounting using sampling techniques are presented.

In Query 2, four variables are created: `mp` of type `Measurement_Point` (line 4); `name` that collects the values of the `id` attribute of each `Measurement_Point` (line 5); `technique` of type `Sampling_Technique` that collects the techniques associated to `Measurement_Points` through the object property `hasSampleTechnique` (line 6); and `type` that collects the types of `Measurement_Points` through `type` attribute (line 7). The variables selected for output are `name`, `technique` and `type` (line 2). Figure 6 shows the obtained values of this query.

– Which are the type and name of each MP, and the corresponding sampling techniques in use;
*Input of Query 2:*

```
1  PREFIX ...
2  SELECT ? name ? technique ? type
3  WHERE {
4  ?mp rdf: type sm: Measurement_Point .
5  ?mp sm:id ? name .
6  ?mp sm: hasSampleTechnique ? technique .
7  ?mp sm: type ? type
8  }
```

| name | technique | type |
|---|---|---|
| "Point4"@ | MUST | "core"@ |
| "Point2"@ | SystT | "Border Router"@ |
| "Point1"@ | SystC | "Border Router"@ |
| "Point3"@ | RandC | "core"@ |
| "Point5"@ | LP | "Border Router"@ |

**Fig. 6.** Output of Query 2

Other example listing the characteristics of a specific MP (MP3) - Query 3 and Figure 7, is provided below.

– Listing the characteristics of a particular MP
*Input of Query 3:*

```
PREFIX ...
SELECT ? name ? cpuDescription ? cpu ? memory
WHERE {
?mp rdf: type sm: Measurement_Point .
?mp sm:id ? name .
FILTER REGEX (? name ," Point3 ").
?mp sm: hasMeasurementNode ?mn.
?mn sm: cpuDescription ? cpuDescription .
?mn sm: currentMemoryLoad ? memory .
?mn sm: currentCpuLoad ?cpu
}
```

| name | cpuDescription | cpu | memory |
|---|---|---|---|
| "Point3"@ | "ARM V7 900MHZ"@ | "18.26"^^<http://v | "86163"^^<http://v |

**Fig. 7.** Output of Query 3

As a final note, once the semantic layer defined as proposed is applied to all nodes in the network under management, monitoring the entire network would be simplified. The provision of an interoperable layer (common ontology), would benefit the development of network monitoring solutions in an objective, systematic and automatic manner.

## 6 Conclusions

This work has proposed the construction of a semantic model to assist the development of solutions based on expert agents for context-aware network monitoring. That is why an ontology has been proposed to describe the domain and a software architecture has been described that provides these high added-value services for network monitoring in a highly automated way. The proposed system was also validated by verifying the competency questions identified and the support of a crucial monitoring task - *traffic accounting*. The experimental validation of the system in a realistic network environment is planned as future work. The authors hope to develop a collaborative system that can anticipate the needs of the system with rules deduced from the behaviour of the system under realistic loading conditions.

**Acknowledgments** - This work has been supported by COMPETE: POCI-01-0145-FEDER-007043 and FCT *Fundação para a Ciência e Tecnologia* within the Project Scope: UID/CEC/00319/2013.

## References

1. J. M. C. Silva, P. Carvalho, and S. R. Lima, "Inside packet sampling techniques: exploring modularity to enhance network measurements," *International Journal of Communication Systems*, vol. 30, no. 6, 2017.
2. M. Grüninger and M. Fox, "Methodology for the Design and Evaluation of Ontologies," in *IJCAI'95, Workshop on Basic Ontological Issues in Knowledge Sharing, April 13, 1995*, 1995. [Online]. Available: http://citeseer.ist.psu.edu/grninger95methodology.html

3. R. Lin, O. Li, Q. Li, and K. Dai, "Exploiting adaptive packet-sampling measurements for multimedia traffic classification," *Journal of Communications*, vol. 9, no. 12, 2014.
4. D. Tammaro, S. Valenti, D. Rossi, and A. Pescapé, "Exploiting packet-sampling measurements for traffic characterization and classification," *International Journal of Network Management*, vol. 22, no. 6, pp. 451–476, 2012.
5. T. Zseby, T. Hirsch, and B. Claise, "Packet sampling for flow accounting: Challenges and limitations," in *International Conference on Passive and Active Network Measurement*. Springer, 2008, pp. 61–71.
6. C. Hu, S. Wang, J. Tian, B. Liu, Y. Cheng, and Y. Chen, "Accurate and efficient traffic monitoring using adaptive non-linear sampling method," in *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*. IEEE, 2008, pp. 26–30.
7. A. N. Mahmood, J. Hu, Z. Tari, and C. Leckie, "Critical infrastructure protection: Resource efficient sampling to improve detection of less frequent patterns in network traffic," *Journal of Network and Computer Applications*, vol. 33, no. 4, pp. 491–502, 2010.
8. Y. Gu, L. Breslau, N. Duffield, and S. Sen, "On passive one-way loss measurements using sampled flow statistics," in *INFOCOM 2009, IEEE*. IEEE, 2009, pp. 2946–2950.
9. S. Yoon, T. Ha, S. Kim, and H. Lim, "Scalable traffic sampling using centrality measure on SDNs," *IEEE Communications Magazine*, vol. 55, no. 7, pp. 43–49, 2017.
10. J.-H. Jun, C.-W. Ahn, and S.-H. Kim, "Ddos attack detection by using packet sampling and flow features," in *proceedings of the 29th annual ACM symposium on applied computing*. ACM, 2014, pp. 711–712.
11. N. Duffield *et al.*, "Sampling for passive internet measurement: A review," *Statistical Science*, vol. 19, no. 3, pp. 472–498, 2004.
12. A. Martinez, M. Yannuzzi, V. López, D. López, W. Ramírez, R. Serral-Gracià, X. Masip-Bruin, M. Maciejewski, and J. Altmann, "Network management challenges and trends in multi-layer and multi-vendor settings for carrier-grade networks," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 4, pp. 2207–2230, 2014.
13. A. K. Y. Wong, P. Ray, N. Parameswaran, and J. Strassner, "Ontology mapping for the interoperability problem in network management," *IEEE Journal on selected areas in Communications*, vol. 23, no. 10, pp. 2058–2068, 2005.
14. "an ontology-based information extraction system for bridging the configuration gap in hybrid sdn environments."
15. H. Xu and D. Xiao, "Applying semantic web services to automate network management," in *Industrial Electronics and Applications, 2007. ICIEA 2007. 2nd IEEE Conference on*. IEEE, 2007, pp. 461–466.
16. C. Rodrigues, S. R. Lima, L. M. Á. Sabucedo, and P. Carvalho, "An ontology for managing network services quality," *Expert Systems with App.*, vol. 39, no. 9, pp. 7938–7946, 2012.
17. P. S. Moraes, L. N. Sampaio, J. A. Monteiro, and M. Portnoi, "Mononto: A domain ontology for network monitoring and recommendation for advanced internet applications users," in *Network Operations and Management Symposium Workshops, 2008. NOMS Workshops 2008. IEEE*. IEEE, 2008, pp. 116–123.
18. A. Simmonds, P. Sandilands, and L. Van Ekert, "An ontology for network security attacks," in *Asian Applied Computing Conference*. Springer, 2004, pp. 317–323.
19. D. V. Silva and G. R. Rafael, "Ontologies for network security and future challenges," in *International Conference on Cyber Warfare and Security*. Academic Conferences International Limited, 2017, p. 541.
20. J. M. C. Silva, P. Carvalho, and S. R. Lima, "Computational weight of network traffic sampling techniques," in *2014 IEEE Symposium on Computers and Communications (ISCC)*. IEEE, 2014, pp. 1–6.