

Arquitetura para gerenciamento de dispositivos através de assistentes virtuais comandados por voz - Resumo Estendido - CTDSI/CTCCSI 2023

Honoré Cesário¹, Gustavo Girão¹, Roger Immich¹

¹ Instituto Metrópole Digital
Universidade Federal do Rio Grande do Norte (UFRN)
honore.cesario@ufrn.br, {girao, roger}@imd.ufrn.br

O gerenciamento, monitoramento e estruturação das redes domésticas tem se tornado cada vez mais complexa. Um dos motivos para isto é devido ao grande aumento no número de dispositivos e equipamentos conectados à sua infraestrutura. Sua configuração tende a se tornar uma tarefa mais comum nos próximos anos visto que, de acordo com pesquisa apresentada por [Titara 2021], 80% das pessoas de classe média pretendem tornar sua casa mais conectada. Destaca-se também que, além da crescente popularização das redes residenciais [Rodrigues et al. 2019], existe também a tendência de haver dispositivos com cada vez mais recursos [Immich et al. 2019].

A grande variedade de fabricantes e a falta de padronização nas funcionalidades disponibilizadas, muitas vezes, apresentam uma interface completamente diferente de um modelo para outro [Fiorenza et al. 2021]. Isso pode levar os usuários a efetuar configurações equivocadas, prejudicando o gerenciamento desses equipamentos. Realizar essas configurações pode ser ainda mais desafiador para pessoas que apresentem algum tipo de deficiência. Dados da última pesquisa realizada pelo IBGE (Instituto Brasileiro de Geografia e Estatística), mostram que mais de 45 milhões de pessoas (23,9% da população) possuem algum tipo de deficiência, onde 36 milhões relatam ter algum tipo de deficiência visual [IBGE 2010]. Esses dados reforçam a necessidade de novas propostas na literatura que promovam um ambiente inclusivo para pessoas que apresentem esse tipo de deficiência.

Os trabalhos relacionados a esse estudo são apresentados em dois grupos: abordagens com monitoramento e acionamento de dispositivos por voz e/ou texto; e trabalhos baseados em linguagem de intenções para configurações de dispositivos de rede. Relativo ao monitoramento e acionamento de dispositivos, há propostas que utilizam assistentes virtuais na função de interface conversacional. Entre estes, podemos citar como por exemplo a adoção do Google Home Assistant [Lago et al. 2021] e da Amazon Alexa [Austerjost et al. 2018]. A inclusão destes dispositivos busca possibilitar que a interação com o usuário ocorra através de comandos por voz de forma natural. *Smartphones* ou dispositivos do tipo Raspberry são apresentados por [Alexakis et al. 2019, Kadali et al. 2020], como dispositivos de entrada para recepção dos comandos, utilizando (em sua maioria) *chatbots* para que ocorra a interação com o usuário. Esses trabalhos, por apresentarem funcionalidades menos complexas, têm como característica a não utilização de uma camada de abstração, dispensando o uso de uma linguagem de intenções.

No grupo que aborda a configuração de dispositivos, um processo de refinamento de intenções utilizando *chatbots*, baseado no *feedback* dos operadores de rede é proposto

por [Jacobs et al. 2018], onde os autores pretendem colaborar para o desenvolvimento de redes autônomas. O *framework* P4I/O, foi desenvolvido por [Riftadi and Kuipers 2019], para configuração de *switches* e tem seu uso voltado a redes baseadas em intenções. O protótipo apresentado por [Alsudais and Keller 2017], fornece uma camada intermediária de comunicação entre diferentes sistemas de gerenciamento de redes e o usuário utilizando linguagem natural. O *framework* Lumi, desenvolvido por [Jacobs et al. 2019], permite que operadores de rede possam expressar suas intenções para configuração de dispositivos utilizando linguagem natural.

Nos trabalhos voltados ao monitoramento e acionamento de dispositivos, percebeu-se a importância na escolha de um extrator de conceitos que possua uma maior abrangência nas linguagens suportadas e que apresente uma base de dados mais ampla, identificou-se também a necessidade de algum mecanismo que garanta uma usabilidade mínima das funcionalidades para os casos de falta de conexão com a Internet, além da possibilidade de implementação de algum mecanismo de autenticação para restringir o acesso à utilização dos serviços [Fernandes et al. 2020]. Já os trabalhos analisados com foco na configuração de dispositivos exigem um grau de conhecimento técnico elevado, utilizam uma base de dados pequena (com poucas sentenças de treinamento) e contam com o *feedback* do usuário para aprimoramento da assertividade no reconhecimento das intenções. Com isso, prejudicam a usabilidade da solução, limitando seu uso apenas para usuários detentores de um maior conhecimento técnico.

De forma a preencher algumas das lacunas mencionadas, o presente trabalho propõe uma arquitetura em camadas, baseada em módulos independentes [Cesário et al. 2022]. A arquitetura faz uso do reconhecimento de comandos em linguagem natural, de forma textual ou por voz, em conjunto com uma linguagem de intenções. Pretende com isso, prover facilidade na configuração de dispositivos de diversos fabricantes, através de uma camada de abstração, permitindo que o usuário realize configurações sem necessariamente ter conhecimento da arquitetura, especificidades das interfaces de gerenciamento e nomenclaturas das funcionalidades, por exemplo, que diferem em cada fabricante. Garantindo com isso comodidade ao usuário e provendo acessibilidade.

A arquitetura proposta mostrou-se viável após os testes realizados com o protótipo que foi desenvolvido. O *chatbot* ao receber solicitações, por voz ou texto (em linguagem natural), identifica a intenção e converte para a linguagem de configuração do dispositivo de rede, em seguida enviando o comando correspondente para execução, tudo isso de forma transparente para o usuário.

A avaliação da proposta foi desenvolvida em dois cenários, através de comandos por texto e de comandos por voz. No cenário onde os comandos foram realizados de forma textual, a ferramenta apresentou uma assertividade média superior a 83%. No segundo cenário, através de comandos por voz, a assertividade média foi de 71,75%. A ferramenta apresentou excelentes resultados, sendo capaz de suportar vários comandos, como por exemplo: configuração de regras de bloqueio de acesso de dispositivos à Internet ou portas de comunicação específicas; e bloqueio de domínios. Sugestões para trabalhos futuros: i) adicionar a arquitetura uma camada para descoberta e registro automático de novos dispositivos que sejam conectados à rede; ii) possibilitar que o *chatbot*, caso questionado, consiga prover informações sobre o *status* dos dispositivos ao usuário.

Referências

- Alexakis, G., Panagiotakis, S., Fragkakis, A., Markakis, E., and Vassilakis, K. (2019). Control of smart home operations using natural language processing, voice recognition and iot technologies in a multi-tier architecture. *Designs*, 3(3).
- Alsudais, A. and Keller, E. (2017). Hey network, can you understand me? In *2017 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pages 193–198.
- Austerjost, J., Porr, M., Riedel, N., Geier, D., Becker, T., Scheper, T., Marquard, D., Lindner, P., and Beutel, S. (2018). Introducing a virtual assistant to the lab: A voice user interface for the intuitive control of laboratory instruments. *SLAS Technology*, 23:476 – 482.
- Cesário, H., Girao, G., Riker, A., Dalmazo, B., and Immich, R. (2022). Arquitetura para gerenciamento de dispositivos através de assistentes virtuais comandados por voz. In *Anais do VI Workshop de Computação Urbana*, pages 1–14, Porto Alegre, RS, Brasil. SBC.
- Fernandes, R., Paz, G., Kretuz, D., Mansilha, R., Jenuario, T., and Immich, R. (2020). S3as: uma solução de autenticação e autorização através de aplicativos de smartphones. *Revista Eletrônica Argentina-Brasil de Tecnologias da Informação e da Comunicação*, 3(1).
- Fiorenza, M., Kretuz, D., Mansilha, R., Macedo, D., Feitosa, E., and Immich, R. (2021). Representação e aplicação de políticas de segurança em firewalls de redes híbridas. In *Anais do XXXIX Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos*, pages 490–503, Porto Alegre, RS, Brasil. SBC.
- IBGE (2010). Censo demográfico 2010. Disponível em: https://biblioteca.ibge.gov.br/visualizacao/periodicos/94/cd_2010_religiao_deficiencia.pdf. Acesso em: 25 de jun. de 2021.
- Immich, R., Villas, L., Bittencourt, L., and Madeira, E. (2019). Multi-tier edge-to-cloud architecture for adaptive video delivery. In *2019 7th International Conference on Future Internet of Things and Cloud (FiCloud)*, pages 23–30.
- Jacobs, A., Pfitscher, R. J., Ferreira, R., and Granville, L. (2018). Refining network intents for self-driving networks. *Proceedings of the Afternoon Workshop on Self-Driving Networks*.
- Jacobs, A. S., Pfitscher, R. J., Ribeiro, R. H., Ferreira, R. A., Granville, L. Z., and Rao, S. G. (2019). Deploying natural language intents with lumi. In *Proceedings of the ACM SIGCOMM 2019 Conference Posters and Demos*, SIGCOMM Posters and Demos '19, page 82–84, New York, NY, USA. Association for Computing Machinery.
- Kadali, B., Prasad, N., Kudav, P., and Deshpande, M. (2020). Home automation using chatbot and voice assistant. *ITM Web of Conferences*, 32:01002.
- Lago, A. S., Dias, J. P., and Ferreira, H. S. (2021). Managing non-trivial internet-of-things systems with conversational assistants: A prototype and a feasibility experiment. *Journal of Computational Science*, 51:101324.
- Riftadi, M. and Kuipers, F. (2019). P4i/o: Intent-based networking with p4. In *2019 IEEE Conference on Network Softwareization (NetSoft)*, pages 438–443.
- Rodrigues, D. O., Santos, F. A., Filho, G. P. R., Akabane, A. T., Cabral, R., Immich, R., Jr., W. L., da Cunha, F. D., Guidoni, D. L., Silva, T. H., do Rosário, D., Cerqueira, E., Loureiro, A. A. F., and Villas, L. A. (2019). Computação urbana da teoria à prática: Fundamentos, aplicações e desafios. *CoRR*, abs/1912.05662.
- Titara, D. (2021). Casa conectada é tendência mundial. *Revista Eletrolar News*, 22(143):54–70.