SBSeg24
SÃO JOSÉ DOS CAMPOS

Identificação de Serviços e Dispositivos em Dados de Motores de Busca para o Enriquecimento de Análise de Vulnerabilidades

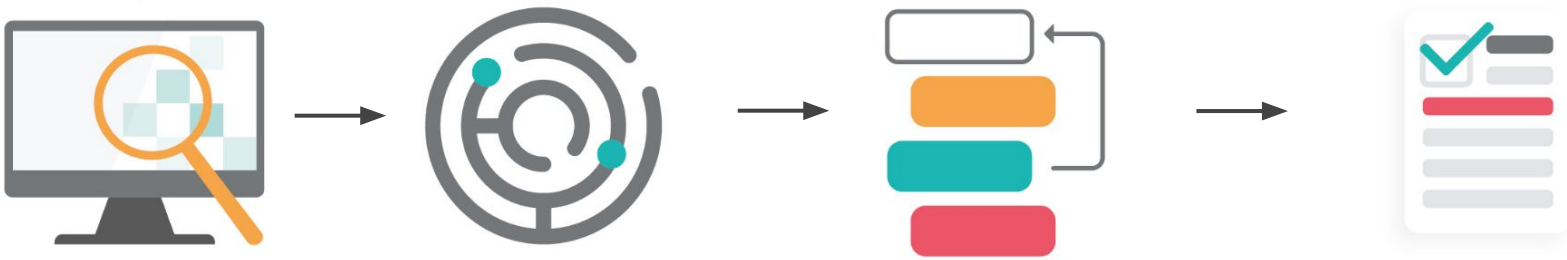
Lucas M. Ponce[1] Indra Ribeiro[1] Etelvina Oliveira[1]
Ítalo Cunha[1] Cristine Hoepers[2] Klaus Steding-Jessen[2]
Marcelo H. P. C. Chaves[2] Dorgival Guedes[1] Wagner Meira Jr.[1]

[1] DCC/UFMG

[2] CERT.br/NIC.br

Introdução
Análise de Vulnerabilidades

Processo contínuo



Descoberta de ativos

Escaneamento de vulnerabilidades

Avaliação de riscos

Relatório das análises e remediação

Motores de busca (de dispositivos)

Introdução
Motores de busca (de dispositivos conectados)

+ 134

Protocols and Products

- amqp
- screenshot
- afp
- airplay
- android_debug_bridge
- ...
- vnc
- windows_exporter
- xiaomi_miio
- yeelight

// PROPERTY

mongodb

Property Name	Type	Required
authentication	boolean	Yes
buildInfo	object	
listDatabases	object	
serverStatus	object	

```
HTTP/1.0 200 OK
Connection: close
Content-Type: text/plain
Content-Length: 85

MongoDB Server Information
{ "process": "mongod", "pid": 9197, ... }
```

Fingerprint


```
{
    "authentication": false,
    "serverStatus": {
        "process": "mongod",
        "pid": 9197, ...
       },
    ...,
    "listDatabases": {
        "totalSize": 274432,
        "ok": 1,
        "databases": [
            {
                "sizeOnDisk": 241664,
                "name": "FlaskLogin",
                "empty": false
            },
            {
                "sizeOnDisk": 32768,
                "name": "Warning",
                "empty": false
            }
        ]
    }
}
```

Motivação e Objetivo

Estatísticas do CPE para o ano de 2023.

2023

Month	New Entries	New Vendors	New Products
January	18,296	419	1,097
February	16,482	276	2,910
March	22,070	324	1,358
April	14,262	221	1,404
May	17,590	266	2,519
June	19,858	260	3,043
July	18,926	272	2,127
August	20,942	240	4,929
September	16,827	239	1,058
October	21,335	340	1,516
November	18,294	341	1,481
December	18,122	299	1,320

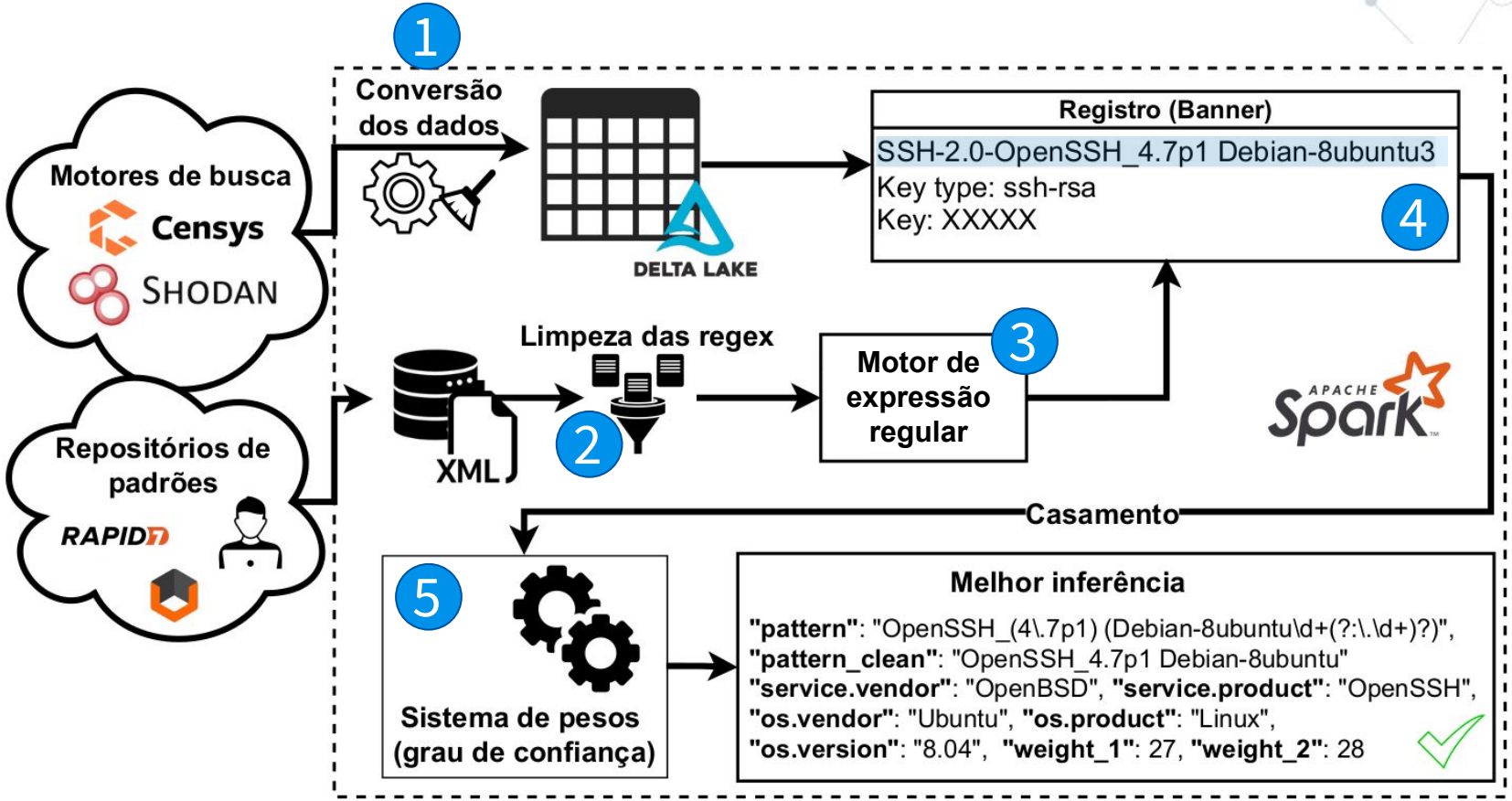
24.762 novos produtos só em 2023 !

Fonte: *https://nvd.nist.gov/products/cpe/statistics*

Objetivo:

- Arcabouço para o processamento **eficiente** de *fingerprints*;
- Aplicável em dados de motores de busca ou de outras fontes.

Arquitetura Geral

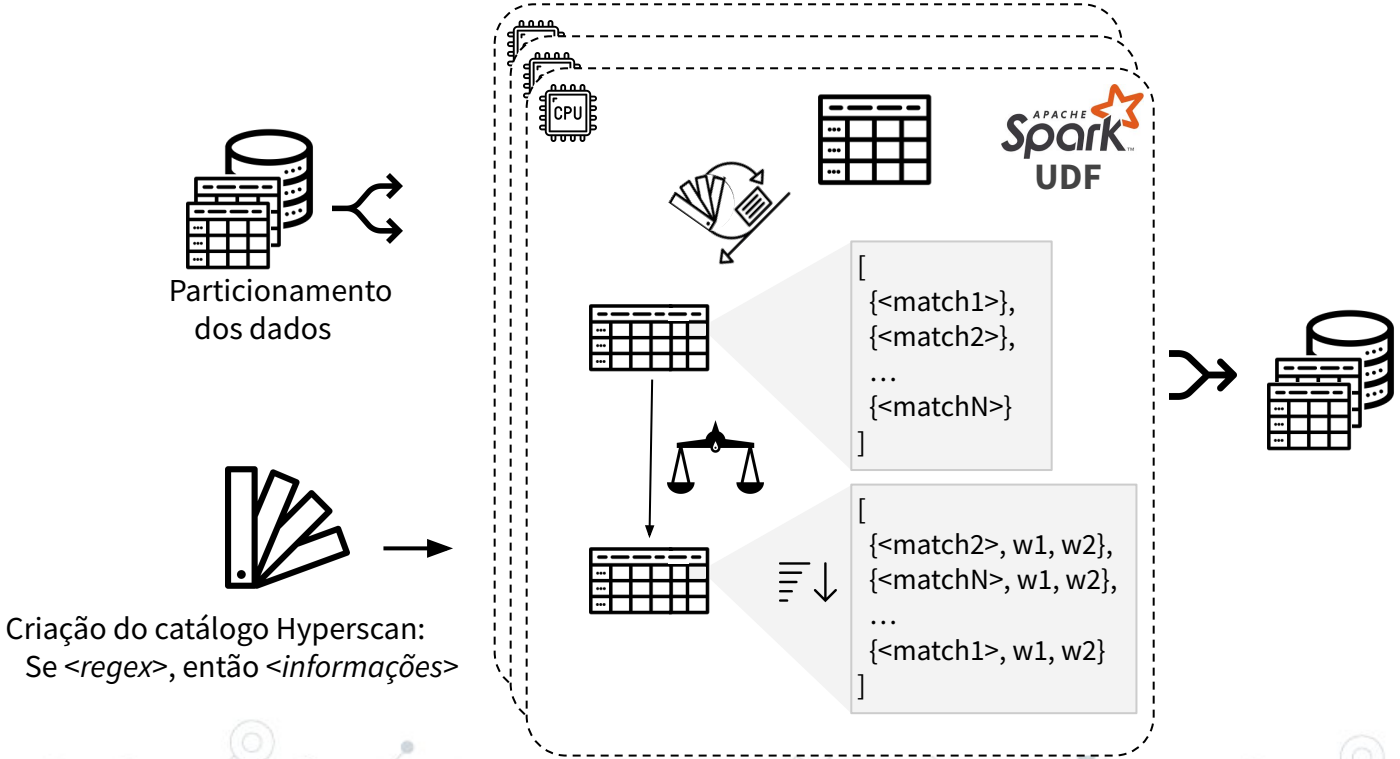


Arquitetura Geral
Coleta e limpeza das regex


```
<fingerprint pattern="OpenSSH_(3\.8\.1p1) (Debian-11ubuntu\d+(?:\.\d+)?)">
    <param name="service.vendor" value="OpenBSD"/>
    <param name="service.family" value="OpenSSH"/>
    <param name="service.product" value="OpenSSH"/>
    <param name="os.vendor" value="Ubuntu"/>
    <param name="os.family" value="Linux"/>
    <param name="os.product" value="Linux"/>
    <param name="os.version" value="4.10"/>
    <param name="os.cpe23" value="cpe:/o:canonical:ubuntu_linux:4.10"/>
</fingerprint>
```

13 categorias de serviços
(HTTP Servers, SSH, etc) = 2432 padrões

Arquitetura Geral
Processamento das regex



Intel Hyperscan, *https://github.com/intel/hyperscan*

Arquitetura Geral
Sistema de peso

pattern: `OpenSSH_(3\.8\.1p1) (Debian-11ubuntu\d+(?:\.\d+)?)`

Filtragem dos termos constantes

pattern_clean: `OpenSSH_3.8.1p1 Debian-11ubuntu`

Resposta do serviço:

```
SSH-2.0-OpenSSH_3.8.1p1 Debian-11ubuntu3

Key type: ssh-rsa
Key: AAAAB3NzaC1yc2...
Fingerprint: 6a:ee:72:19:72:18:e9: ...

Kex Algorithms:
        curve25519-sha256
        curve25519-sha256@libssh.org
        ecdh-sha2-nistp521
        ecdh-sha2-nistp384
...
```

```
weight1 = length(fingerprint)
        = length("OpenSSH_3.8.1p1 Debian-11ubuntu3") = 32
```

```
weight2 = SequenceMatcher(fingerprint, pattern_clean)
weight2 = SequenceMatcher(
            "OpenSSH_3.8.1p Debian-11ubuntu",
            "OpenSSH_3.8.1p1 Debian-11ubuntu3"
   ) = 14 + 16 = 30
```

93.5% do fingerprint é formado por termos obrigatórios

Mais forte do que o padrão: *OpenSSH_.+*
weight1: 32, weight2: 7 (21,8%)

Avaliação Experimental

Estudo avaliativo

Caracterização inicial (1º semestre de 2021 e 2023)

Tabela 1. Relação do número de registros e IPs com a porcentagem de casamentos.

Métrica	Registros	# IPs
Total	330.414.756	19.557.988
% Inferidos	59,8	68,9

175 min (no total) ou 28 seg/dia !

Ambiente: Spark com 10 VCPU e 20 GB RAM

Tabela 2. Relação dos padrões por categoria de serviço.

Fonte	Registros (%)	IPs (%)	Padrões	
			(#)	(%)
HTTP (Servers)	41,34	36,25	402	90,33
DNS	9,15	25,01	57	81,42
HTML	8,53	19,61	195	43,33
SSH (Recog)	7,08	27,16	129	86
HTTP (Auth)	5,41	14,64	60	80
Apache (OS)	5,18	7,23	27	71,05
Telnet	4,32	18,63	57	39,58
HTTP (Cookies)	1,95	1,6	62	76,54
SNMP (SysDescr)	1,67	2,99	133	23,58
SMB (OS)	1,18	2,23	60	78,94
FTP	0,75	1,26	81	54,72
SSH (Outros)	0,25	0,51	65	98,48
SNMP (SysObjId)	0,02	0,12	7	16,66
SMB (LM)	0,01	0,05	5	62,5
NTP	< 0,01	< 0,01	1	1,33

Estudo avaliativo

Relação dos padrões encontrados

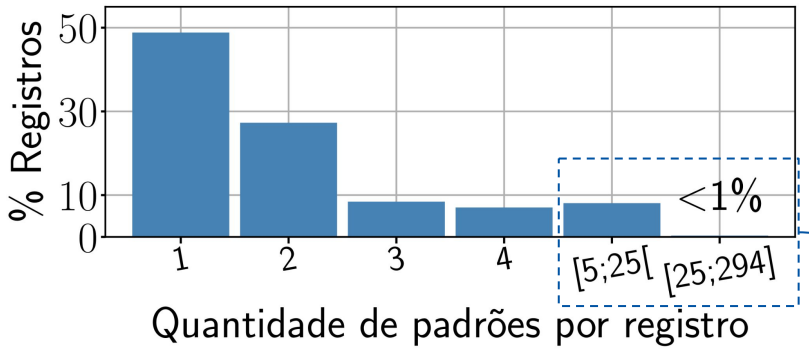


Figura 3: Histograma da quantidade de padrões por registro.

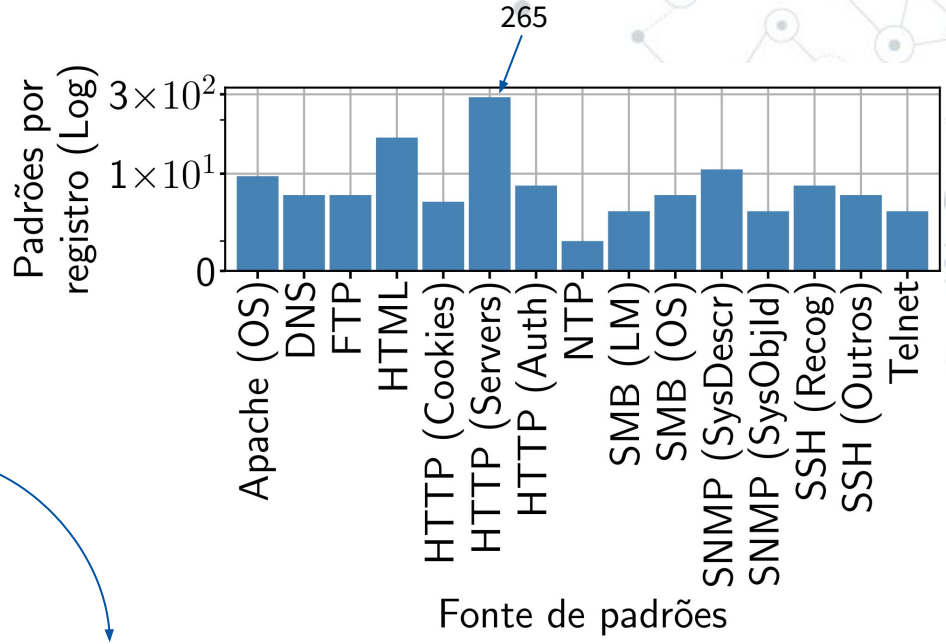


Figura 4: Número máximo de padrões de uma mesma fonte/registro.

Múltiplos **tipos** de serviços
ou
Múltiplos **padrões** sobre um mesmo serviço

Exemplo: o registro de 20 mil caracteres com diversos rastros de serviços *web*, impressoras, Windows e até serviços Linux.

Estudo avaliativo

Identificação de informações sobre o sistema operacional e o *hardware*

Tabela 5: Inferências sobre o S.O. e o *hardware*.

Shodan	Inferido	Sistema Operacional		*Hardware*	
		% IPs	% Registros	% IPs	% Registros
✗	✗	69,08	75,96	80,16	91,57
✗	✓	11,90	18,21	18,43	7,81
✓	✗	2,70	1,77	0,92	0,32
✓	✓	16,32	4,06	0,49	0,30

1,6x mais IPs

14,1x mais IPs

Poucos casos onde apenas o Shodan infere

```
RTSP/1.0 401 Unauthorized
CSeq: 1
WWW-Authenticate: Digest realm="Hikvision",\
 nonce="3a359…", stale="FALSE"
WWW-Authenticate: Basic realm="/"
```

Cameradar, *https://github.com/Ullaakut/cameradar*

Casos de uso

Caso de uso 1

Revisitando o census de 2016 sobre o protocolo SSH na Internet

Census: 15.646.188 IPs no mundo
2023: 2.957.615 IPs no Brasil (18,9% do Census global)
3,2x mais que 2021

Tabela 6: Top 5 produtos SSH em 2023.

Produto	% IPs	Censys 2016
DropBear	77,7	2º (13,8 %)
OpenSSH	16,9	1º (74,4 %)
ROSSSH	2,3	4º (2,9 %)
Ausente	0,7	7º (0,3 %)
Comware	0,5	12º (0,2 %)

1) Versões de 2019 em 93,8 %
2) Existem versões desde 2003 !

Diminuição do SSH-1.X

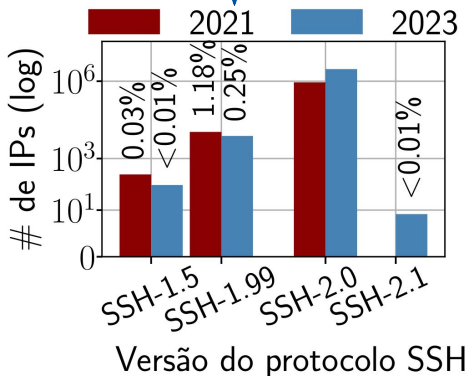


Figura 5: Diversidade do protocolo SSH no Brasil em 2023.

Internet Census, 2016. https://t.ly/aQRM2

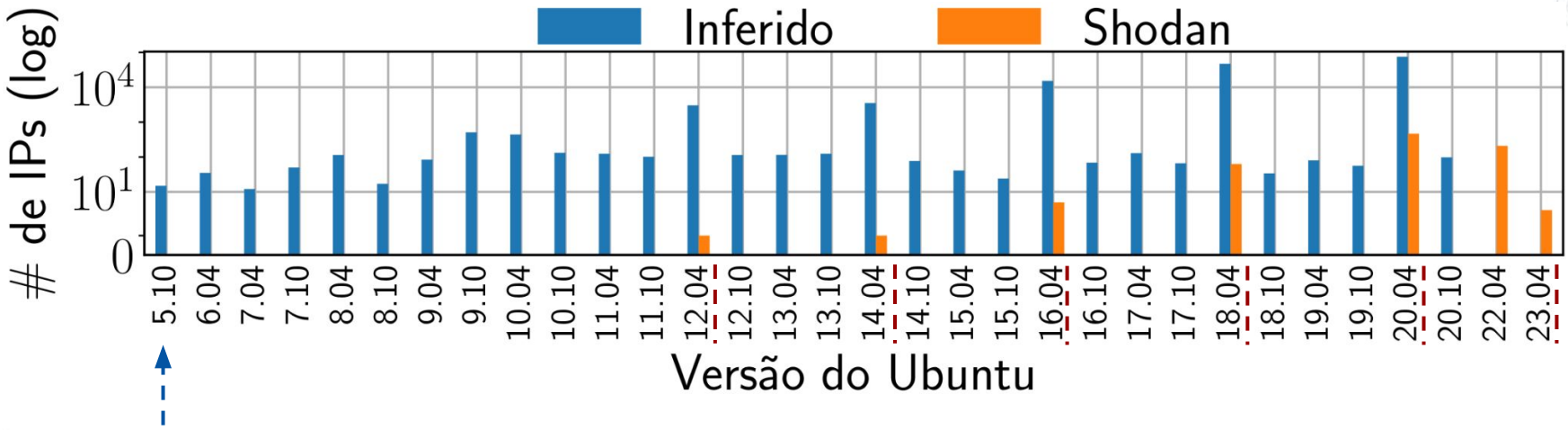
Caso de uso 2

Utilização de sistemas Ubuntu desatualizados

Shodan: 198 mil IPs, apenas 749 com versão

421 mil IPs

147 mil !



Fim do suporte estendido em 2007.

Figura 6: Quantidade de dispositivos com o sistema Ubuntu no ano de 2023.

Considerações finais

Arcabouço para **processar padrões em larga escala** para o reconhecimento de aplicações e dispositivos conectados à Internet:

- Enriquecimento de dados provenientes de motores de busca.

Comparação experimental com as inferências nativas do Shodan:

- Identificação de mais serviços: **sistema operacional** (em 1,6 vezes mais IPs) e informações sobre **dispositivos** (14,1 vezes);
- Qualidade das informações é melhor estruturada;

Disponível em: *https://github.com/lucasmsp/tlhop-library*

Trabalhos futuros

- Análise:
  - Comparação de vulnerabilidades entre abordagens (Shodan vs enriquecida);
  - Identificação de dispositivos em IPs dinâmicos.

- Arcabouço:
  - Inclusão de novas fontes e novos padrões de *fingerprints*;
  - Avaliação de outras abordagens para o ranqueamento dos casamentos obtidos;

Projeto TLHOP/SAM

UFMG CERT.br

Lucas M. Ponce

Indra Ribeiro

Etelvina Oliveira

Ítalo Cunha

Cristine Hoepers

Klaus Steding-Jessen

Marcelo H. P. C. Chaves

Dorgival Guedes

Wagner Meira Jr.

SBSeg24
SÃO JOSÉ DOS CAMPOS

Identificação de Serviços e Dispositivos em Dados de Motores de Busca para o Enriquecimento de Análise de Vulnerabilidades

Lucas M. Ponce[1] Indra Ribeiro[1] Etelvina Oliveira[1]
Ítalo Cunha[1] Cristine Hoepers[2] Klaus Steding-Jessen[2]
Marcelo H. P. C. Chaves[2] Dorgival Guedes[1] Wagner Meira Jr.[1]

[1] DCC/UFMG

[2] CERT.br/NIC.br