Analisando conceitos de Segurança da Informação na Internet do ponto de vista de um Autonomous System(AS) utilizando calculadora categorial

Germano de Mello Andersson

Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)
Caixa Postal 15.064 – 91.501-970 – Porto Alegre – RS - Brasil

Resumo. *Com o advento da computação em nuvem e a facilidade de acesso à Internet de nossa década, as organizações tem percebido a necessidade de disponibilizar seus aplicativos na rede mundial de computadores com o objetivo de proporcionar aos seus usuários e clientes acesso a seus sistemas, onde quer que eles estejam. Neste contexto, a adoção de boas práticas que aumentem a segurança da informação é fundamental. Esta abordagem visa mostrar como a aplicação de conceitos de teoria das categorias pode ser extremamente poderoso para detecção de evidências como, por exemplo, um ponto único de falha de uma determinada rede.*

1. Teoria das Categorias

Teoria das categorias estuda objetos e setas entre eles[Asperti e Longo 1991]. Inicialmente, qualquer abordagem a respeito da semântica destes objetos e setas não são relevantes, e esta é uma das características que tornam esta teoria tão poderosa. Quando descemos um nível de abstração e atribuimos significados aos elementos dos objetos e suas relações, algumas propriedades e cálculos verificados na teoria poderão trazer resultados interessantes no contexto desta implementação.

1.1 Categoria

Uma categoria é constituída de entidades primitivas denominadas objetos e morfismos entre objetos [Menezes and Haeusler 2001]. A seguir, é apresentando o conceito de categoria como uma seis-upla:

$C = \langle ObC, MorC, \delta 0, \delta 1, i, o \rangle$, onde:

Obc é uma coleção de Objetos.

Morc é uma coleção de Morfismos.

$\delta 0, \delta 1$: Morc → Obc são operações de Origem e Domínio respectivamente.

i: Obc → Morc é uma operação de Identidade.

o: $(Morc)^2$ → Morc é uma operação parcial de Composição de morfismos.

1.2 Categoria Dual

Para um dada uma categoria C, a categoria dual de C será obtida com a inversão de seus morfismos, ou seja:

$C_{op} = \langle ObC, MorC, \delta 1, \delta 0, i, o_{op} \rangle$

1.3 Categoria Livremente Gerada à partir de um grafo

Uma categoria livremente gerada à partir de um grafo é o resultado do enriquecimento do grafo com as operações identidade e composição.

1.4 Diagrama
Para uma dada categoria C, um diagrama em C é uma multicoleção de objetos e de morfismos de C tal que, para cada morfismo do diagrama, os seus correspondentes objetos origem e destino são elementos do diagrama[Menezes and Haeusler 2001].

1.5 Isomorfismo
Um morfismo f: A → B de uma dada categoria C é um isomorfismo se e somente se ele possui um morfismo inverso g: B → A tal que:

$$g \circ f = iA \; ; \; f \circ g = iB$$

1.6 Cone
Para uma dada categoria C, um cone de um diagrama D consiste em um objeto K juntamente com uma multicoleção de morfismos desse objeto para quaisquer objetos de D satisfazendo a seguinte condição de comutatividade:

Para quaisquer objetos A,B de D temos que:

para qualquer ka: K → A e qualquer kb: K → B, se existir um morfismo ab: A → B então kb = ab o ka.

1.7 Cocone
Cocone é o dual de Cone.

1.8 Produto Fibrado
Produto Fibrado é o limite (cone ótimo) de um diagrama composto por dois morfimos com um mesmo objeto destino. Através de dois morfismos f: A → C e g: B → C, o Produto Fibrado pode ser calculado através do produto entre A e B equalizado por f e g.

1.9 Soma Amalgamada
Soma Amalgamada é o dual de Produto Fibrado.

2. Redes de Computadores
Uma série de computadores interconectados[Tanenbaum 1997].

2.1 Conectividade entre redes
Conectividade entre redes é a capacidade de uma rede alcançar outra rede. É comum que duas redes estejam conectadas por mais de um caminho.

2.2 Autonomous System (AS)
Cada instituição que administra uma ou mais redes presentes na tabela de roteamento externo é denominada de AS(Autonomous System). Caso uma instituição tenha necessidade de políticas de roteamento distintas, esta instituição deverá ser mais de um AS. Idealmente, um AS deveria ter apenas um super prefixo de rede que englobasse todos sub prefixos de rede necessários, porém na prática tal planejamento e previsionabilidade se tornou inviável.

2.3 Autonomous System Number (ASN)
ASN é um identificador do tipo inteiro que representa, de forma única, um AS. No Brasil temos 1364 ASNs e em Porto Alegre 23, com dados atualizados em 24 de novembro de 2011.

2.4 AS neighbor
Um AS neighbor (vizinho) é um AS que está exatamente a um salto de distância, ou seja, diretamente conectado.

2.5 AS_PATH
AS_PATH é um atributo de um protocolo de roteamento externo. Ele pode ser visto como uma estrutura de dados pilha que contém todos saltos necessários para conectividade entre duas redes por um determinado caminho. Cada salto é representado pelo ASN do AS correspondente.

2.6 Atribuição de Endereçamento de Rede
Com o objetivo de organizar a distribuição de endereçamentos de rede na internet, foi criado um cartório da internet (IR), organizado em uma hierarquia de 3 níveis[RFC 2050]. No topo, temos a IANA, que tem a responsabilidade sob todo espaço de endereçamento. A IANA distribui então blocos de endereçamento para os cartórios regionais, que operam em um espaço geopolítico envolvendo diversas nações e são os responsáveis por repassar subblocos de endereçamento para os cartórios locais. Enfim, os cartórios locais distribuem endereços de rede para empresas provedoras de internet(ISP) ou para empresas usuário final(EU). Os cartórios locais estão dispostos, na sua maioria, em regiões nacionais.

Do ponto de vista do Brasil, a hierarquia de atribuição de endereçamento de rede é a seguinte: IANA > LACNIC > Registro.br.

2.7 Internet Service Provider (ISP)
Um ISP é um AS que recebe endereços de rede de um cartório local e os distribui para utilização de empresas terceiras.

2.8 End User (EU)
Um EU é um AS que recebe endereços de rede de um cartório local e consome a utilização destes endereços.

2.9 Internet
A internet, ou rede mundial, pode ser definida como um conjunto de redes interconectadas[Tanenbaum 1997]. É uma rede livre de escala que pode ser visualizada como um grafo onde os nodos representam as redes e as arestas a conexão entre elas. O modelo de comunicação de dados utilizado é o TCP/IP, portanto o encaminhamento de tráfego acontece em camada de Rede(Layer3) em equipamentos nominados roteadores. Cada AS é representado por um ou mais roteadores.

3. Segurança da Informação

A informação é o bem mais precioso para grande maioria das empresas. Por isso, segurança da informação deixou de ser uma abordagem apenas das equipes de computação e está sendo tratada como fator chave para o sucesso de um negócio. Existem diversos conceitos que definem a segurança da informação, mas para fins deste artigo utilizaremos apenas três: confidencialidade, integridade e disponibilidade.

3.1 Confidencialidade
Prover confidencialidade da informação é garantir o acesso a informação tão somente às entidades legítimas, ou seja, àquelas autorizadas pelo proprietário da informação.

3.2 Integridade
Prover integridade da informação é garantir que ela, manipulada, mantenha todas as características originais estabelecidas pelo proprietário da informação, incluindo controle de mudanças e garantia do seu ciclo de vida (nascimento,manutenção e destruição).

3.3 Disponibilidade
Prover disponibilidade da informação é garantir que ela esteja sempre disponível para o uso legítimo, ou seja, por aqueles usuários autorizados pelo proprietário da informação.

4. Modelo Categorial
Com base nos conceitos apresentados acima, vamos introduzir a categoria InternetCat e, com algumas operações categóricas, descobrir características interessantes quanto ao comportamento desta rede de computadores.

4.1 Definição
Seja InternetCat a categoria que representa a rede de computadores internet onde:

a) cada objeto é um AS composto por elementos que são suas redes.

b) cada morfismo é um as_path que representa o caminho de conectividade entre redes de dois ASs;

c) a operação identidade (i) representa que, para um determinado AS, todas as suas redes estão conectadas diretamente umas as outras e o morfismo identidade é um as_path no qual todos elementos na pilha são o próprio ASN.

d) a operação composição (o) representa a capacidade de alcançar uma rede que está a mais de um salto de distância, ou seja, não está diretamente conectada nem está em um AS neighbor. Basicamente, esta operação é o empilhamento de dois as_path.

4.2 Estendendo a categoria InternetCat
É importante citar que alguns morfismos gerados pela operação composição são considerados inválidos na implementação de um protocolo de roteamento externo. Para definir estas peculiaridades, vamos estender a definição de objetos e morfismos da categoria InternetCat.

Um objeto AS da categoria InternetCat possui um atributo chamado ASN do tipo inteiro e um atributo chamado classe do tipo string. O primeiro será utilizado para identificação do AS. O segundo será utilizado para informar qual a classe do AS: 'ISP' ou 'EU'.

Um morfismo as_path da categoria InternetCat possui dois atributos booleanos chamados loop e disable_forward. Para podermos gerar resultados interessantes desta categoria, será necessário adicionar duas funções a uma calculadora categorial básica.

4.3 Adicionando duas funções a uma calculadora categorial básica

1) Operação Loop: Esta operação é necessária para que sejam mapeados os morfismos que geram loops na topologia de rede da internet. O algoritmo desta nova operação é o seguinte:

```
void loop() {
   for each  as_path in Morc:
      as_path = false
      for each  AS in Obc:
         // se existe um morfismo com destino AS que contenha AS.ASN e não é
         // identidade, então ele é loop.
         If (     ( δ1(as_path).isequal(AS) ) AND
                  ( as_path.contains(AS.asn) ) AND
                  ! (as_path.isequal(i()) )
         )
         then
                  as_path.loop = true
                  break
}
```

2) Operação disable_forward: ASs End User devem anunciar apenas as rotas das suas redes. Isto evita que tráfegos de encaminhamento atravessem estas redes, ou seja, apenas tráfego com origem ou destino devem alcançá-las. O algoritmo desta nova operação seria o seguinte:

```
void disable_forward() {
   for each  as_path in Morc:
      as_path = false
      for each  AS in Obc:
         // se existe um morfismo com origem de um AS End User que AS.ASN não é o
         // primeiro elemento da  pilha as_path, então ele é  disable_forward.
         If (     ! (as_path.first().isequal(AS.asn) ) AND
                  ( AS.classe='EU' )
         then
                  as_path.disable_forward = true
                  break
}
```

Para fins de clareza, morfismos loop e disable_forward não serão exibidos nos diagramas a seguir.

5. Aplicação de Conceitos Categoriais

O principal objetivo deste documento é apresentar como uma calculadora categorial pode ser utilizada como uma ferramenta para visualização e depuração de problemas de roteamento entre redes. O ápice deste estudo é a utilização dos cálculos de Produto Fibrado e Soma Amalgamada como detectores de ponto único de falha na conectividade de uma determinada rede com a Internet.

5.1 Categoria Livremente Gerada

Um diagrama apresentando um grafo, conforme a figuras 1, pode ser entendido como o estado inicial de conectividade entre ASs, onde há arestas apenas entre ASs e seus vizinhos. Podemos interpretá-lo como um estágio inicial da Internet. A categoria livremente gerada à partir deste grafo é ilustrada nas figuras 2 e 3, e representa o próximo estado de conectividade entre estes ASs, onde cada AS anuncia suas redes conhecidas e, sincronamente, cada AS recebe anúncio de seus vizinhos. Os ciclos seguintes de geração livre da categoria promoverão a topologia da Internet tal como ela é nos dias de hoje, com a devida dinamicidade para reconfigurá-la em função de falha de elementos que comprometam um caminho com várias conectividades.

O anúncio e o aceite de uma rede na troca de informação entre ASs é simétrico[RFC 1930], o que garante o mesmo número de passos para ambos procedimentos e portanto o mesmo número de ciclos de geração livre da categoria para criação de um morfismo conectividade.

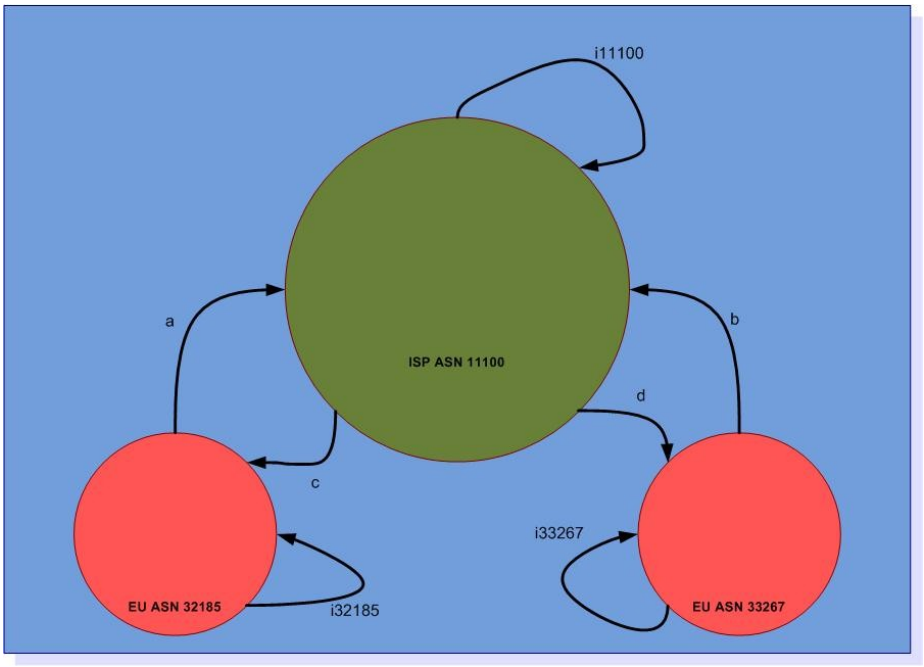

Figura 1 – Diagrama de um grafo.

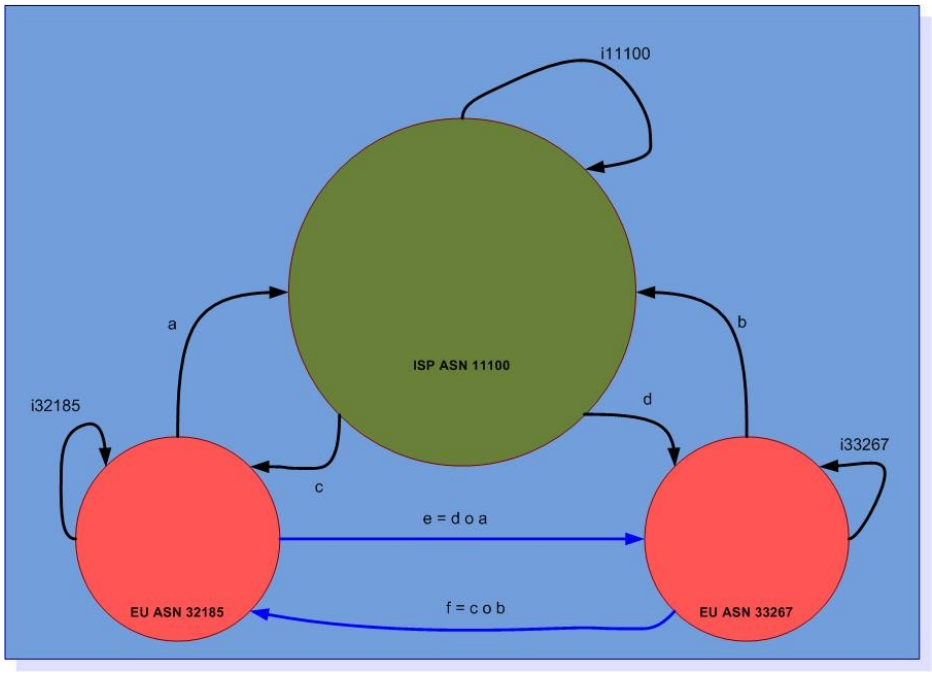

Figura 2 – Diagrama de InternetCat, gerado à partir do grafo da figura 1.

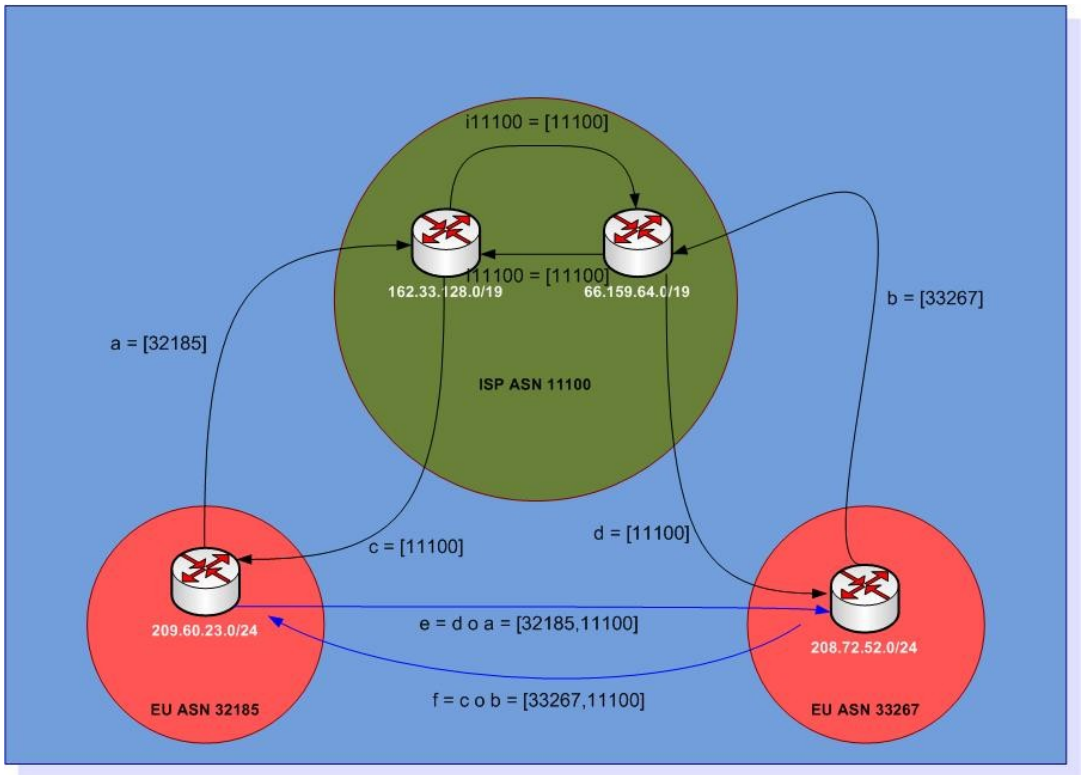

Figura 3 – Diagrama de InternetCat da figura 2, agora com a semântica dos morfismos.

5.2 Isomorfismo

Um isomorfismo pode ser visto como uma rede presente em ASs distintos, conforme figura 4:

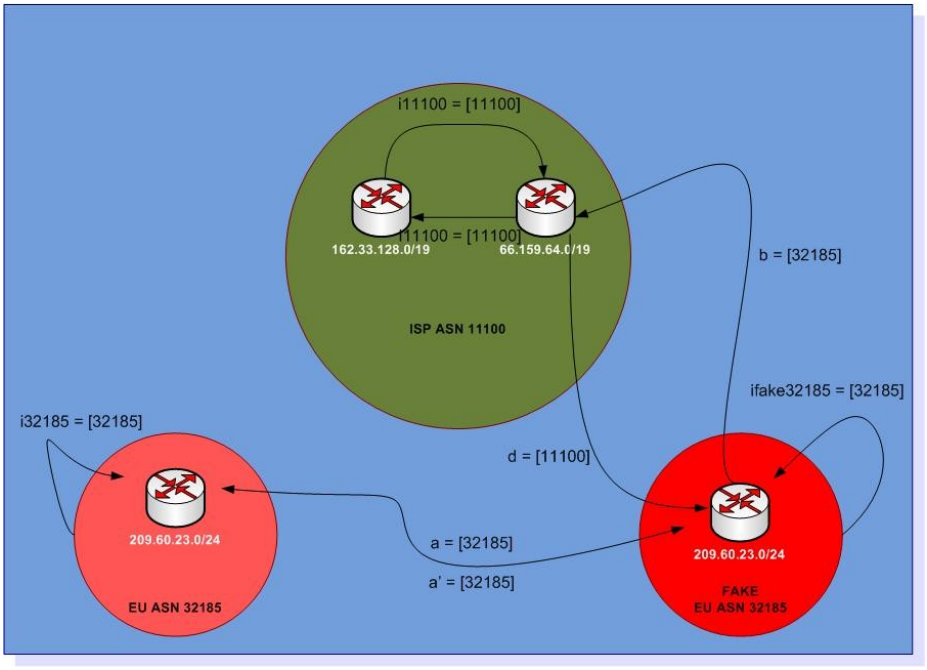

Figura 4 – Diagrama de Internetcat com um isomorfismo.

Por definição, uma rede presente em mais de um objeto é inválido no contexto de redes de computadores. Analisando este cenário, percebe-se que a presença de um isomorfismo representa a detecção de um falso anúncio de rotas. O anúncio falso de rotas é uma técnica famosa de exploração de vulnerabilidade de protocolo de roteamento externo. Bastam alguns anúncios falsos propagados com a 'promessa' de ser o melhor caminho para uma dada conectividade que em poucos minutos todo tráfego para uma rede pode estar passando primeiramente pelo atacante. A exploração desta técnica é classificada como 'man in the middle', e pode fazer com que dados sejam interceptados e analisados, violando os conceitos de confidencialidade e integridade da informação. Para evitar a exploração desta vulnerabilidade o NIST (National Institute os Standards and Technology) lançou documento[NIST 800-54] recomendando melhores práticas a serem adotadas por analistas de infraestrutura de redes.

5.3 Cone

Para as próximas análises, vamos tomar como base o diagrama apresentado na figura 5, que exibe uma topologia onde um AS usuário final, está conectado a internet através de dois ISPs, numa topologia conhecida como Multihoming. Tal configuração é utilizada em redes críticas, com o objetivo de que a falha de um dos seus provedores não afete em nada a disponibilidade de suas aplicações.

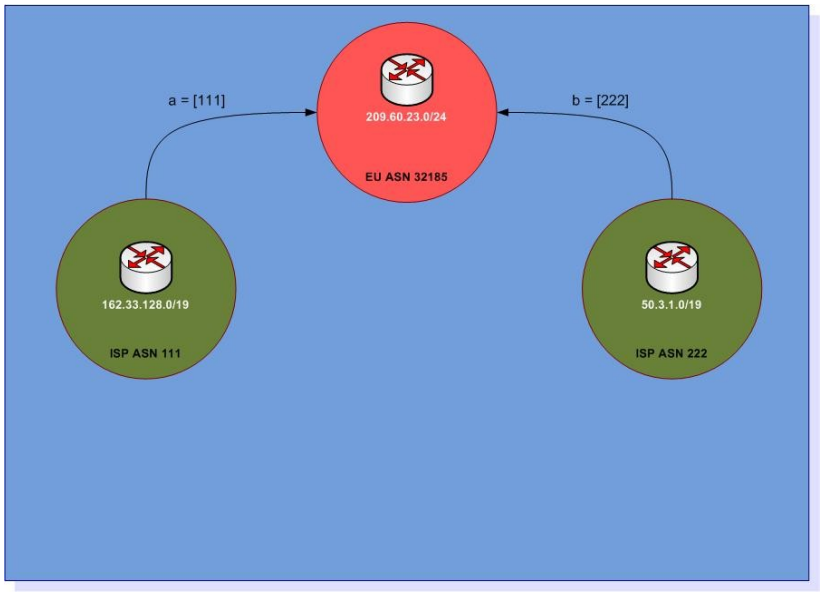

Figura 5 – Diagrama de uma categoria representando topologia multihoming.

Todo objeto vértice de um cone deste diagrama pode ser interpretado como um AS com redes cobertas pela alta disponibilidade oferecida pela topologia Multihoming, ou seja, se uma rede pertence a um AS cone deste diagrama, ela continuará tendo conectividade mesmo em caso de falha de um dos provedores.

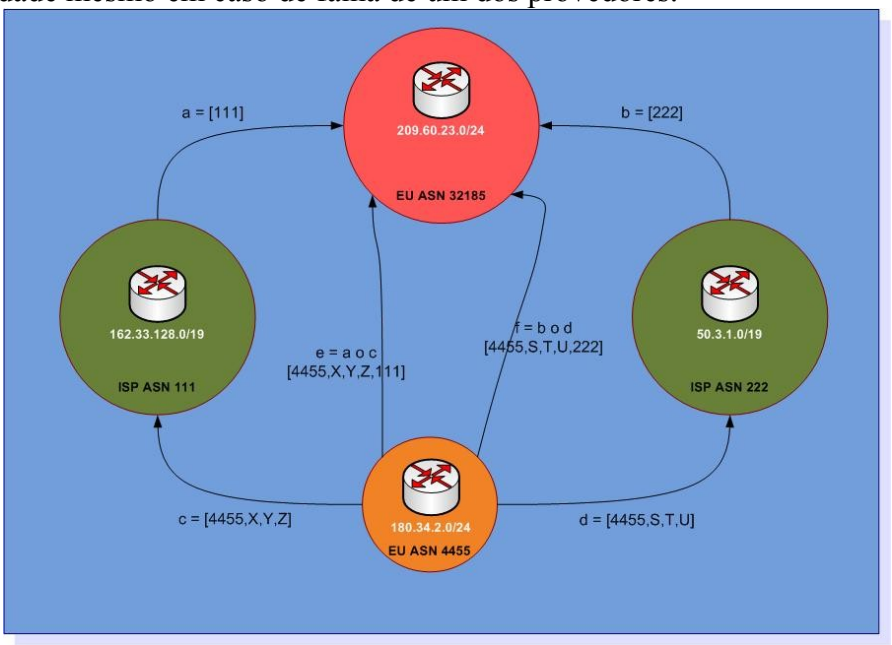

Figura 6 – AS4455 mantém contato com AS32185 mesmo em caso de falha de um dos ISPs.

5.4 Cocone

Aproveitando a persistência de propriedades de uma categoria e sua dual, fica fácil identificar que todo cocone do diagrama da figura 4 pode ser intepretado como um AS com redes balanceáveis, ou seja, o AS usuário final poderá alcançar estas redes tanto por um provedor quanto pelo outro, podendo equilibrar a utilização de banda disponível por

seus provedores.

5.5 Produto Fibrado

Um produto fibrado é um limite, portanto um cone ótimo. O conceito de ótimo traz consigo a idéia de estrutura essencial, sem redundância, e isto pode nos ajudar a calcular um temido elemento em infraestrutura de redes: o ponto único de falha.

Vejamos a figura 7, onde temos 2 cones, sendo um deles produto fibrado, do diagrama apresentado na figura 4.

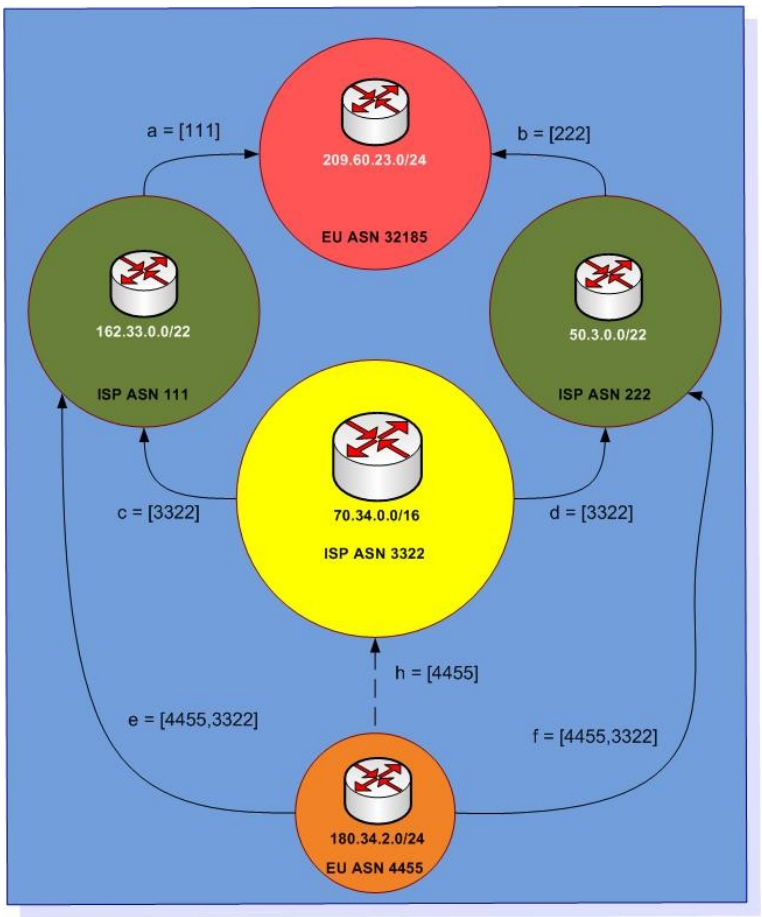

Figura 7 – Cálculo Produto Fibrado mostra que o AS4455 só possui caminho para o AS32185 através do AS3322.

Podemos chegar a conclusão que na existência de um produto fibrado do diagrama que representa um AS e seus provedores, todo AS que for vértice de no máximo 1 cone terá no produto fibrado um ponto único de falha, ou seja, perda da conexão total em caso de falha do AS que exerce papel de produto fibrado. Este estado da rede ilustra em que situação uma configuração multihoming pode estar vulnerável a indisponibilidade, e, por consequência, a falha de uma estratégia de segurança da informação.

5.6 Soma Amalgamada

Por dualidade, podemos concluir que a existência de um AS que é soma amalgamada deste mesmo diagrama representa um ponto único de falha para o alcance de qualquer outro AS que seja vértice de no máximo um cocone.

6. Conclusões

Tais exposições demonstram como um ferramental que aparentemente tem fins acadêmicos pode ser útil para o mercado de tecnologia. De fato, a calculadora categorial fornece resultados extremamente relevantes de modo simples e eficaz.

Segurança da Informação e Redes de Computadores são núcleos críticos no contexto de Infraestrutura Tecnológica. A rápida evolução da internet impulsiona a necessidade de profissionais com excelência em ambas áreas, fazendo com que analistas de infraestrutura cada vez mais precisem conhecer prioritariamente a arquitetura das soluções do que uma tecnologia específica. O estudo de teorias como a teoria das categorias, que forçam o pensamento para o abstrato, são altamente recomendadas para atingir este perfil profissional.

7. Referencias

Menezes, P. B. and Haeusler, E. H. (2001). “Teoria das Categorias para Ciência da Computação”, Série Livros Didáticos. Instituto de Informática da UFRGS.

Asperti, A. and Longo, G. (1991). “Categories, Types, and Structures: An Introduction to Category Theory for the Working Computer Scientist”. The MIT Press.

Tanenbaum, Andrew S. (1997). “Redes de Computadores”, 3ª Edição

Kuhn, R.; Sriram, K.; Montgomery, D. (2007) “Recommendations of the National Institute of Standards and Technology”, NIST Special Publication 800-54

Internet Registry Ip Allocation Guidelines - http://www.ietf.org/rfc/rfc2050.txt

A Border Gateway Protocol 4 (BGP-4) - http://www.ietf.org/rfc/rfc4271.txt

Guidelines for creation, selection, and registration of an Autonomous System (AS) - http://www.ietf.org/rfc/rfc1930.txt

Information Systems Security Association (ISSA) - https://www.issa.org/