

UNIVERSIDADE FEDERAL DE SANTA CATARINA – UFSC
CENTRO SÓCIO ECONÔMICO
DEPARTAMENTO DE ECONOMIA E RELAÇÕES INTERNACIONAIS

DANIEL SOUSA BARRA

**OTIMIZAÇÃO DE PORTFÓLIO PARA UMA CARTEIRA DE CRIPTOMOEDAS: UMA
ABORDAGEM EM REINFORCEMENT LEARNING**

Florianópolis, 2018

DANIEL SOUSA BARRA

**OTIMIZAÇÃO DE PORTFÓLIO PARA UMA CARTEIRA DE CRIPTOMOEDAS:
UMA ABORDAGEM EM REINFORCEMENT LEARNING**

Monografia submetida ao curso de Ciências Econômicas da Universidade Federal de Santa Catarina, como requisito obrigatório para a obtenção do grau de Bacharel em Ciências Econômicas.

Orientador: Prof. Dr. Helberte João França Almeida

FLORIANÓPOLIS

2018

UNIVERSIDADE FEDERAL DE SANTA CATARINA
CURSO DE GRADUAÇÃO EM CIÊNCIAS ECONÔMICAS

A banca examinadora resolveu atribuir a nota 9.0 ao aluno Daniel Sousa Barra na disciplina CNM 7107 – Monografia, pela apresentação deste trabalho.

Banca Examinadora:

Prof. Dr. Helberte João França Almeida
Orientador

Prof. Dr. Jaylson Jair da Silveira
Membro da Banca

Prof. Dr. Maurício Simiano Nunes
Membro da Banca

AGRADECIMENTOS

Primeiro, agradeço ao meu pai, pelas diversas mentorias prestadas ao longo da vida, que me tornou a pessoa e o profissional que sou hoje.

Agradeço à Taynara, minha namorada, por ter aguentado as muitas noites que passei desenvolvendo esse trabalho.

Agradeço também à minha equipe na Navarra, minha empresa, que gerou muitos dos insights que precisei ao longo dessa trajetória, e pela compreensão com a dedicação necessária para a conclusão.

Por último, agradeço ao Helberte, meu orientador, por ter me guiado no desenvolvimento desse trabalho.

RESUMO

Criptomoedas têm sido um assunto muito discutido em todo o mundo a partir do segundo semestre de 2017. Embora muitos economistas discutam a validade das mesmas como unidade de valor, pouco tem sido discutido quanto à sua aplicabilidade como ativo financeiro. A alta volatilidade que esses ativos possuem, atrelada a uma liquidez considerável dentro do espectro do mercado de capitais, permite que se criem novas oportunidades de investimento para investidores que buscam alternativas arrojadas. Diante desse cenário, este estudo busca encontrar um modelo de otimização de portfólio utilizando uma carteira teórica composta por cinco criptomoedas. Para tanto, se utiliza de um algoritmo de otimização para definir, dentro de uma faixa de valores, qual a porcentagem da carteira que deve ser alocada em cada criptomoeda para cada período de rebalanceamento. Ao final de cada período, a análise verifica para quais moedas a direção de alocação (compra ou venda) foi a correta, e então é aplicado um algoritmo de *reinforcement learning* para aumentar a faixa das que tiveram a direção correta, e reduzir a faixa das que tiveram a direção errada, de forma a melhorar as próximas decisões de alocação. Os resultados demonstraram que as carteiras com otimização obtiveram melhor desempenho do que as que tinham porcentagens fixas de alocação (sem rebalanceamento), e também que as carteiras que aplicaram *reinforcement learning*, novamente tiveram melhor desempenho, do que as que utilizam otimização sem *reinforcement learning*.

Palavras-chave: Aprendizado por reforço. Criptomoedas. Otimização de portfólio.

ABSTRACT

Cryptocurrencies have been the subject of a lot of discussion worldwide from the second semester of 2017 onwards. Even though many economists doubt their worth as a store of value, little has been discussed about their worth as a financial asset. The high volatility of these assets, attached to a considerable liquidity inside the spectrum of financial markets, allows for new investment opportunities for investors that look for risky alternatives. Before this scenario, the study aims to find a model of portfolio optimization utilizing a theoretical wallet composed by five cryptocurrencies. To do so, an optimization algorithm is used to define, inside a range of values, which percentage of the wallet has to be allocated to each cryptocurrency for each rebalancing period. At the end of each period, the analysis checks which currencies have been sold or bought correctly, and a reinforcement learning algorithm is applied to raise the range of the coins that had the right direction and reduce the range of the ones with the wrong direction, to enhance the next allocation decisions. The results show that portfolios with optimization performed much better than the ones with fixed allocation percentages (no rebalancing) and also that the portfolios with reinforcement learning performed much better than the ones without it.

Keywords: Reinforcement learning. Cryptocurrencies. Portfolio optimization.

LISTA DE ILUSTRAÇÕES

Figura 1 – O <i>hash pointer</i>	14
Figura 2 – Elos da <i>blockchain</i>	14
Figura 3 – Transações possíveis em um único bloco	15
Figura 4 – Volatilidade dos ativos em um ano.....	17
Figura 5 – Risco e retorno	19
Figura 6 – Funcionamento do <i>reinforcement learning</i>	21
Figura 7 – Carteira teórica versus Bitcoin	27
Figura 8 – Carteira teórica versus as cinco maiores criptomoedas	28
Figura 9 – Retornos diários da carteira teórica versus Bitcoin.....	29
Figura 10 – Resultado com bandas de 40%	30
Figura 11 – Retornos com bandas de 40%	30
Figura 12 – Pesos fixos para cada criptomoeda	31
Figura 13 – Retornos diários com pesos fixos.....	32

LISTA DE TABELAS

Tabela 1 - Percentagem de <i>hashes</i> por país	16
Tabela 2 - Indicadores por carteira	33

SUMÁRIO

1. INTRODUÇÃO	10
1.1 TEMA E PROBLEMA DE PESQUISA	11
1.2 OBJETIVOS	11
1.2.1 Objetivo Geral	11
1.2.2 Objetivos Específicos	11
1.3. JUSTIFICATIVA	11
2. REFERENCIAL TEÓRICO	13
2.1 CRIPTOMOEDAS	13
2.2 OTIMIZAÇÃO DE PORTFÓLIO	18
2.3 APRENDIZADO POR REFORÇO	20
3. ESCOLHA DA CARTEIRA DE CRIPTOMOEDAS BASEADO NO APRENDIZADO POR REFORÇO	23
3.1 SELEÇÃO DAS CRIPTOMOEDAS	23
3.2 DADOS, COMPOSIÇÃO DO PORTFÓLIO E O ALGORITMO DE APRENDIZAGEM POR REFORÇO	24
3.2.1 Implementação Computacional	25
4. RESULTADOS	27
4.1 RESULTADOS DA CARTEIRA	27
4.2 COMPARAÇÕES COM OUTRAS ESTRATÉGIAS	29
5. CONSIDERAÇÕES FINAIS	34
REFERÊNCIAS	36
APÊNDICE	39

1. INTRODUÇÃO

Ao longo do ano de 2017, segundo dados do *site* CoinMarketCap (2018), referência em dados sobre o segmento, o valor de mercado das criptomoedas foi de 500 bilhões de dólares. No cenário de queda na taxa de juros brasileira, a tendência é que os agentes passem a ter mais apetite por investimentos de risco. Desta forma, abre-se uma oportunidade para os investidores adotarem mecanismos de carteiras de ativos para diversificação do risco, mas mantendo a rentabilidade trazida pelos ativos mais voláteis. Há hoje diversos investidores que acreditam no conceito de criptomoedas para o futuro, mas que têm grande incerteza com relação a qual ativo em específico irá prosperar.

Diante disto, o presente estudo busca elaborar uma carteira teórica de criptomoedas. Assim, esta carteira irá oferecer uma alternativa de investimento ao investidor comum. Isso acontecerá por que a carteira representará a performance das principais criptomoedas do mercado, realizando a diversificação de risco através de uma cesta de moedas, isto é, representa uma oportunidade de investimento em criptomoedas com diluição do risco específico de cada criptomoedas individualmente.

Para elaborar esta carteira de moedas, utilizar-se-á, através de algoritmos de otimização, diferentes métodos de rebalanceamento de carteiras e pesos por ativo na carteira. Mais especificamente, irá adotar aprendizagem de máquina (*reinforcement learning*) para escolher os melhores pesos dos ativos no início de cada período. Como é dado de alta frequência (*big data*), o período de rebalanceamento será diário.

Para proteger a integridade da carteira, adicionar-se-á à cesta de moedas apenas as moedas mais respeitadas e com melhor reputação no mercado, aplicando critérios de elegibilidade. Os critérios a serem observados são a capitalização total da criptomoeda, bem como sua capitalização líquida, que é a capitalização que considera a quantidade de criptomoedas para livre circulação, que impacta no seu volume transacionado livremente no mercado.

Portanto, ao realizar este estudo espera-se que a carteira teórica apresente um desempenho melhor que as carteiras com estratégias ingênuas (apenas Bitcoin), ou com pesos fixos (carteiras com 20% de diferentes criptomoedas). Além disso, ao adotar aprendizado de máquinas para encontrar os pesos ideais na carteira, utilizar-se-á os métodos contemporâneos da literatura especializada.

1.1 TEMA E PROBLEMA DE PESQUISA

Dentro do mercado de criptomoedas, há diversos ativos que investidores podem adquirir. Esses ativos, entretanto, são extremamente voláteis, e muitos possuem uma baixa liquidez. A escolha dos ativos certos, e a proporção ideal de cada um na carteira tem influência direta sobre a rentabilidade da carteira como um todo. O tema dessa pesquisa é caracterizado como sistemas adaptativos, pois os agentes podem evoluir e utilizar diferentes estratégias a cada período para alcançar a melhor alocação ótima de uma carteira de forma a maximizar os retornos. Desta forma, serão utilizados algoritmos de otimização e aprendizados de máquinas, tanto para definição dos pesos de cada moeda, como do limite máximo e mínimo de alocação por moeda.

1.2 OBJETIVOS

A seguir apresentam-se os objetivos geral e específicos deste trabalho.

1.2.1 Objetivo Geral

Compreender as técnicas de otimização de portfólios e aprendizado com reforço e, baseado nelas, desenvolver uma carteira hipotética de criptomoedas e avaliar a sua rentabilidade.

1.2.2 Objetivos Específicos

Os objetivos específicos para cumprir o objetivo geral são três:

- Apresentar os principais aspectos presentes na literatura a respeito de criptomoedas, otimização de portfólio e *reinforcement learning*;
- Desenvolver uma metodologia de alocação de carteiras baseada em aprendizado por máquinas, assim como demonstrar a sua implementação computacional;
- Comparar a rentabilidade do modelo escolhido com os principais ativos do mercado de criptomoedas e com carteiras que utilizam outras estratégias de alocação.

1.3. JUSTIFICATIVA

O mercado financeiro é um dos mais concorridos ambientes de negócios do mundo. É também um mercado repleto de recursos e com demanda de evolução constante. Com a

infraestrutura de big data e um algoritmo de otimização, pode-se fazer testes conclusivos sobre os resultados gerados com parametrização de escolha de ativos no mercado de criptomoedas. Isso pode gerar um retorno financeiro significativo, visto que implementará estratégias pouco convencionais com um modelo de execução diferente de tudo que existe hoje no mercado. Assim, pode-se gerar retornos bastante acima do mercado atual, inclusive para mercados institucionais, e com um risco de operação reduzido e mensurável.

A motivação para a realização deste trabalho encontra-se nos estudos de Choi et al (2009), no qual os autores observam que investidores individuais exageram em suas conjecturas com base em resultados pessoais de períodos passados. Investidores que têm retornos maiores nos seus fundos aumentam a taxa de poupança mais do que investidores com experiências piores. Essa atitude é análoga ao comportamento de *reinforcement learning*, que recompensa os acertos. Assim, os autores destacam a possibilidade de modelar decisões de investimento individuais através de *reinforcement learning*.

Athey (2018) afirma que o *machine learning* faz parte do futuro das ciências econômicas. Segundo a autora, um dos usos do *reinforcement learning* em específico é a avaliação de políticas públicas – no caso, buscando a alocação mais eficiente dos recursos dispensados pelo governo em políticas de segurança e bem-estar. Por sua vez, Mullainathan e Spiess (2017) destacam a possibilidade de utilizar o *reinforcement learning* aplicado a estudos econométricos, e mostram o reconhecimento de padrões em bancos de dados de grande escala como uma função importante do aprendizado por máquinas.

A presente monografia encontra-se estruturada, além dessa introdução, por mais quatro capítulos. O capítulo dois apresenta o referencial teórico, que abordará os princípios no qual se baseia o trabalho, trazendo as principais contribuições sobre os temas discutidos. Mais especificamente, apresenta a literatura de criptomoedas, otimização de portfólio e aprendizado por máquinas. O capítulo três mostra a metodologia empreendida, bem como o modelo e o algoritmo de aprendizado por reforço adotado. Por sua vez, o capítulo quatro traz os resultados e uma apresentação e discussão dos mesmos. Por fim, o capítulo cinco traz as considerações finais do trabalho.

2. REFERENCIAL TEÓRICO

2.1 CRIPTOMOEDAS

O uso de itens simbólicos de meios de troca é uma tradição que vem já dos nossos ancestrais. A humanidade já passou por diversos ativos que serviam como tal, desde metais preciosos até dólares. No momento atual, ocorre uma nova disrupção nesse sentido, se trata do advento das criptomoedas. Apesar da ideia de uma moeda virtual ser bem mais antiga, foi a partir da criação de Satoshi Nakamoto, o Bitcoin, que as criptomoedas efetivamente passaram a estar acessíveis ao público amplo. Hoje ela possui valor de mercado de 142 bilhões de dólares.

Nakamoto (2009) define uma moeda virtual como uma corrente de assinaturas digitais, nas quais dois agentes transferem recursos por meio da validação de uma chave pública e uma chave privada, ambas criptografadas. Toda a ideologia por trás da criação das criptomoedas gira ao redor de ter um sistema financeiro que não dependa da confiança em instituições, e que preserve a privacidade dos agentes que transacionam.

Segundo Chiu e Koepl (2017), em um cenário em que a moeda é virtual, os meios de pagamentos não são nada além de um conjunto de bits. Mais especificamente, se tratam de grandes volumes de dados criptografados que compõem uma transação, ou conjunto de transações, que depois compõem um chamado bloco. Quando um bloco é validado, passa a fazer parte de uma corrente (daí provém o nome *blockchain*).

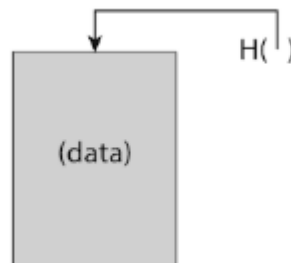
Em moedas do sistema financeiro tradicional, as chamadas *fiats*, organizações como bancos centrais controlam a emissão, o estoque, e realizam controles antifraude, o que torna as barreiras para falsificar essas moedas alta, mas ainda longe do impossível (NARAYANAN et al, 2016). Nas criptomoedas, não há autoridade central, e sim há diversos agentes que em conjunto compõem um consenso sobre quais transações são válidas e quais não são. São eles os mineradores, que servem como árbitros remunerados para validar as assinaturas digitais, e também impedir que algum agente do sistema trapaceie. No caso das criptomoedas, os rígidos controles de segurança são assegurados pela tecnologia por trás do sistema.

O primeiro conceito chave das criptomoedas está nos chamados *hash pointers*. Ainda segundo Narayanan et al (2016), uma função *hash* é uma função matemática que obedece a três propriedades, a saber: i) O *input* pode ser uma variável de texto de qualquer tamanho; ii) O output tem tamanho fixo; e iii) É computável, ou seja, que dada uma variável de texto, seja possível descobrir o output daquela função em um tempo razoável.

Um *hash pointer* é um ponteiro, que permanentemente aponta para um conjunto de

dados, e dessa forma vai de um lugar para outro, como mostra a Figura 1.

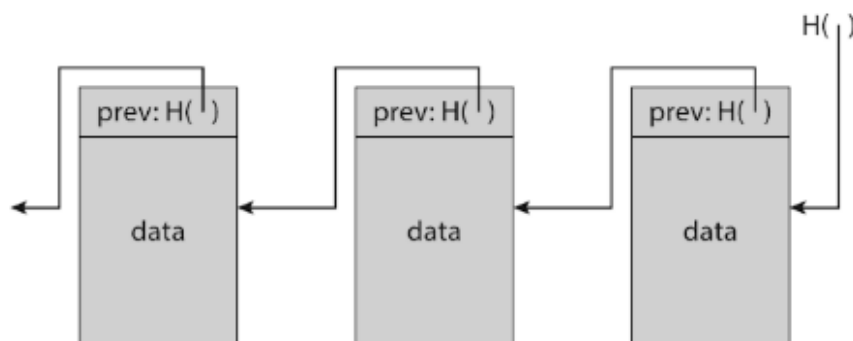
Figura 1– O *hash pointer*



Elaborado por Narayanan et al (2016).

Sua função primordial é ligar os elos da corrente de blocos, conforme a Figura 2 mostra.

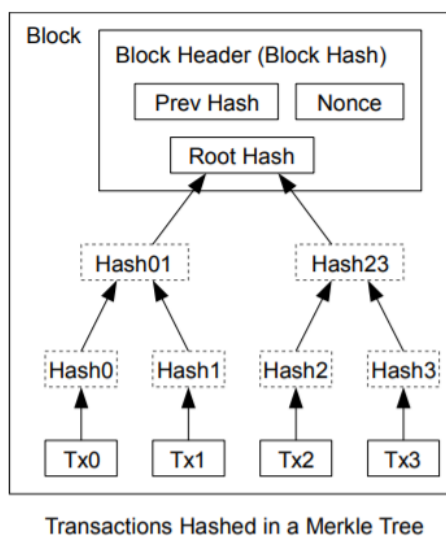
Figura 2– Elos da *blockchain*



Elaborado por Narayanan et al (2016).

Este conjunto sequencial de dados explica a dificuldade de alterar o histórico de transações já realizadas, pois seriam necessários mudar todos os elos desde o início para que fosse modificada a corrente sem os “árbitros” perceberem. Dentro de cada um desses blocos, há um conjunto de transações, que são mantidas unidas por meio de uma árvore de Merkel, na qual um *hash pointer* pode apontar para blocos de dados separados, que apontam para outros blocos, e assim se ramifica a gama de transações possível em um único bloco. A Figura 3, retirada de Nakamoto (2009) demonstra esse efeito.

Figura 3—Transações possíveis em um único bloco



Elaborado por Nakamoto (2009).

Pode-se observar pela imagem que em um bloco, há, além das transações, um *hash* principal do bloco, na qual estão contidos um *hash pointer* que direciona para o bloco anterior, e também um conjunto de números que compõe um problema matemático que os mineradores competem para resolver.

Os blocos se tornam válidos a partir do instante em que um dos mineradores encontra a solução para o problema matemático (atividade que se chama minerar o Bitcoin). Esse minerador então envia esse bloco resolvido para todos os outros mineradores. Nesse instante há a primeira confirmação de um agente sobre aquele bloco. Caso os outros mineradores concordem com a validade daquele bloco, geram mais confirmações, e a corrente vai se tornando cada vez mais forte, e menos propensa a seguir em outra ramificação como a oficial, até eventualmente ser o alvo do *hash pointer* do próximo bloco. A corrente mais longa é sempre considerada a correta pelo ecossistema de mineradores (NARAYANAN et al., 2016).

É de total interesse dos mineradores ser o primeiro a encontrar, pois dessa forma ele é recompensado com um pagamento em novas moedas emitidas junto ao bloco. Segundo Nakamoto (2009), isso tem dois motivos. Primeiro como um incentivo para que os agentes minerem a criptomoeda, e também para colocar as moedas em circulação, visto que não há uma autoridade central emitindo as moedas. Outro tipo de incentivo que os mineradores possuem são as taxas de transferência de quem envia moeda para outra carteira paga.

Neste mercado, a maior parte dos mineradores está na Ásia. A Tabela 1 mostra como se dá a distribuição por país do poder computacional por trás do Bitcoin.

Tabela 1- Percentagem de *hashes* por país

País	<i>Hashes</i>
China	78,0%
República Tcheca	9,9%
Índia	1,1%
Estados Unidos	1,0%
Islândia	0,6%
Israel	0,3%
Rússia	0,2%
Outros	8,9%

Elaboração própria com base em dados de J.P. Morgan.

Há diversas outras criptomoedas além do Bitcoin, e estas utilizam outras metodologias, mas sempre girando em torno do mesmo conceito de corrente de blocos (*blockchain*). Segundo a base de dados CoinMarketCap (2018), em março de 2018 haviam 1592 criptomoedas catalogadas, movimentando cerca de 25 bilhões de dólares por dia em negociações, e com um valor de mercado total de mais de 400 bilhões de dólares. Como comparativo, isso é mais que o dobro do volume da Bovespa no mesmo período. As 5 moedas com maior capitalização de mercado são, de maior para menor são: Bitcoin, Ether (ETH), Ripple (XRP), Bitcoin Cash (BCH) e Litecoin (LTC). Somadas, possuem um total de 230 bilhões de dólares.

Ether, segundo o site oficial, é uma moeda criada para remunerar a plataforma descentralizada Ethereum, que roda contratos inteligentes em uma *blockchain* proprietária, permitindo inúmeras aplicações que necessitem de consenso. É o segundo maior ativo em valor de mercado, com 53 bilhões de dólares.

Ripple, segundo o site oficial, é uma moeda direcionada para facilitar as transações financeiras globais, via *blockchain* própria, que conecta bancos, meios de pagamento, bolsas de criptomoedas, e outros para fornecer transferências com o mínimo de atrito possível e total segurança. Já possuem grandes empresas do setor financeiro tradicional como parceiros, como Santander, American Express, e outros. Tem um valor de mercado de cerca de 27 bilhões de dólares.

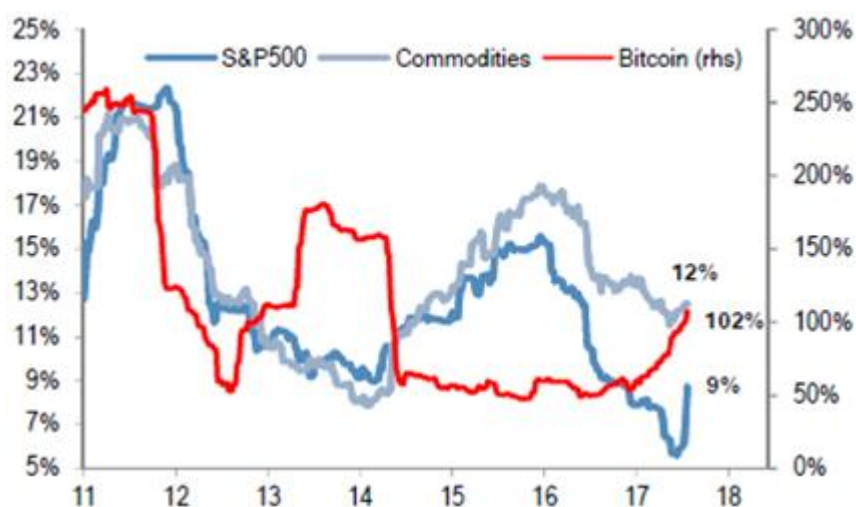
Bitcoin Cash, segundo o site oficial e o *site* Ledger Wallet, surgiu de um rompimento de parte da comunidade com o Bitcoin tradicional, no qual a corrente rompeu e seguiu em duas direções (o chamado *fork*), propondo solução para alguns dos problemas do Bitcoin tradicional, principalmente no que se refere ao uso da moeda como meio de pagamento. Um

dos principais pontos de divergência em relação ao Bitcoin tradicional é que os blocos do Bitcoin Cash são 8 vezes maiores. Possui movimentações diárias de cerca de 400 milhões de dólares.

O Litecoin é uma moeda virtual que tem como objetivo levar o custo das transações para próximo de zero e mais rápido do que as outras criptomoedas. Além de um volume de negociação diária de cerca de 425 milhões de dólares por dia, ocupa o quinto lugar em valor de mercado, com cerca de 8,7 bilhões de dólares.

Uma característica específica desse tipo de ativos é a alta volatilidade. Pode-se observar na Figura 4 como essa volatilidade evoluiu no período de um ano, de fevereiro de 2017 até fevereiro de 2018, em relação ao S&P 500 (principal índice de ações norte-americanas) e as *commodities*. A volatilidade foi dez vezes maior do que a dos outros dois ativos.

Figura 4– Volatilidade dos ativos em um ano



Elaborado por J.P. Morgan.

Segundo Yermack (2013), a correlação do Bitcoin com outras moedas é próxima de zero, e isso praticamente elimina o uso dela como moeda de troca. “Bitcoin imposes large risks on its owners, because it has excessive volatility and fails to exhibit correlation with the behavior of other currencies”.(YERMACK, 2013, p. 7). De acordo com o autor, isso torna a moeda mais parecida como um ativo especulativo do que um meio de troca.

2.2 OTIMIZAÇÃO DE PORTFÓLIO

Há um ditado popular que diz que “não devemos colocar todos nossos ovos na mesma cesta”. Esta frase refere, nesse contexto, ao processo de diversificação de uma carteira de investimentos, de forma a eliminar o risco diversificável.

Segundo Santos (2010), há dois tipos de riscos que afetam uma carteira de investimentos. São eles o risco diversificável e não diversificável (sistêmico). O segundo caso refere-se ao risco a que todos os títulos e valores mobiliários estão expostos, tais como flutuações macroeconômicas, incerteza política e outros. Já o risco diversificável se trata dos riscos específicos ao ativo, e que, por meio de um bom balanceamento de portfólio, podem ser minimizados.

Cada ativo, por si só, possui um retorno esperado e um nível de risco intrínseco individual. Em conjunto, entretanto, os ativos que compõe a carteira geram ao conjunto um novo nível de risco e retorno esperado global. O que deixa a pergunta: como escolher quais ativos e em que proporção para obter o melhor resultado? Esse resultado pode ser medido pelo risco, pelo retorno, ou por ambos.

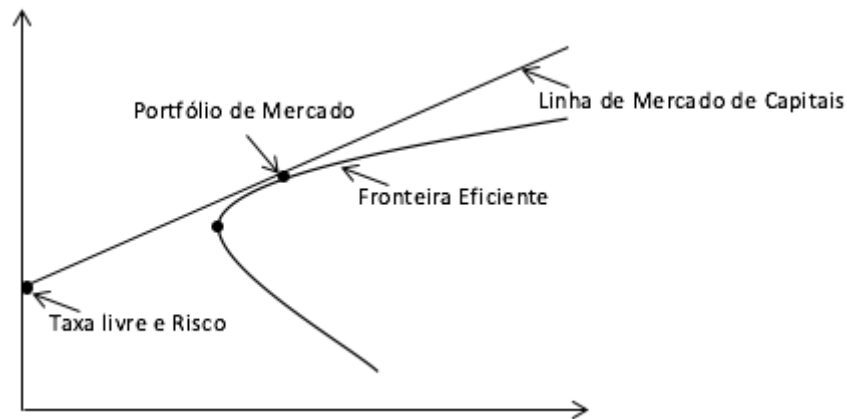
Segundo Boyd et al (2017), otimização de portfólio é o processo de responder à pergunta de quais seriam os melhores parâmetros a serem escolhidos para uma alocação de carteira, que maximizem ou minimizem algum resultado. O mais comum são as otimizações que objetivam maximizar o retorno e minimizar o risco. É o caso do mais famoso modelo de otimização de portfólio, desenvolvido por Markowitz (1952). Importante ressaltar que otimizações sempre dão resultados do que teria sido o melhor resultado, e não do que realmente será melhor no futuro. “We are really asking what would have been the best constant allocation, had we known future returns” (BOYD et al, 2017, p. 21)

Segundo Krink e Paterlini (2008), desde o trabalho de Markowitz em 1952, o estudo de otimização de portfólio tem sido foco de diversos trabalhos acadêmicos e aplicações no mercado financeiro. Essa metodologia usa a abordagem da média-variância, que objetiva minimizar o risco (dado pela variância dos retornos passados) e maximizar o retorno (dado pela média dos retornos passados). Entretanto, segundo o autor, esse modelo é simplista por não utilizar medidas de risco mais realistas e limitações do mundo real.

De acordo com Pereira e Henrique (2016), o modelo de Markowitz é um modelo de programação quadrática que possibilita gerar portfólios para níveis esperados de risco ou de retorno. Ao se realizar o processo para vários níveis de retorno, cria-se a denominada curva de Markowitz que determina a fronteira para a qual as diferentes combinações de proporções de ativos de uma carteira promovem os maiores retornos com os menores riscos possíveis

(MARKOWITZ, 1952).

Figura 5– Risco e retorno



Elaborado por Castellano (2017).

Markowitz (1952) define que para que a seleção de portfólio, é necessário que se cumpra ao menos duas condições. Primeiro, que o investidor acate as limitações de risco e retorno. Segundo, que uma medida de retorno esperado e covariância entre ativos seja possível.

De acordo com Pereira e Henrique (2016), a equação que se usa nesse modelo para o cálculo de risco (variância, tratada pela letra V) e retorno (retorno esperado, dado pela letra E) são:

$$E = \sum_{i=1}^n X_i * \mu_i \quad (1)$$

$$V = \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} * X_i * X_j, \quad (2)$$

no qual X_i é a proporção da carteira alocada no ativo i-ésimo; μ_i é o valor esperado para o retorno do ativo i-ésimo (dado pela média dos retornos passados); σ_{ij} é a covariância do ativo i em relação ao ativo j (No caso de serem iguais, então será o desvio padrão do ativo); e X_j é a proporção da carteira alocada no ativo j-ésimo.

Segundo Santos (2016), tal modelo supõe que uma função utilidade dos riscos e retornos da carteira pode expressar as preferências do investidor. Tendo em vista que os ativos possuem um comportamento do qual o retorno tem correlação positiva com o risco, encontrar uma solução que ao mesmo tempo melhore o retorno e diminua o risco é uma atividade complexa. Portanto, em geral, quanto mais avesso ao risco um determinado investidor for, menor a rentabilidade excessiva (em relação ao ativo livre de risco) este deve demandar.

As principais críticas a esse modelo são, primeiro, que ela considera os retornos de

ativos financeiros como tendo distribuição normal, que não se sustenta no mundo real. Segundo ele não considera a necessidade de inserir restrições não lineares. Um exemplo dessas restrições é o lote mínimo que pode ser negociado de um determinado ativo.

De acordo com Kitschke (2014), um modelo que está sendo bastante discutido para otimização de portfólios é o SLSQP (*Sequential Least-Squares Quadratic Programming*). Ainda segundo o autor, o modelo consiste em um otimizador que define uma direção e um valor de variação nos parâmetros para cada vez que roda um teste, até que encontre um valor que seja o melhor da série, e que fazendo novas modificações não resulte em um resultado superior. O grande benefício dessa função é que ela possui a capacidade de incluir tanto as restrições de igualdade ou desigualdade (por exemplo, que os pesos dos ativos da carteira devem somar 100%), quanto limites de bandas para cada parâmetro (por exemplo, que os pesos individuais nunca superem 40%).

2.3 APRENDIZADO POR REFORÇO

Reinforcement learning, ou aprendizado por reforço, segundo Murphy (1998) se trata de um problema de maximização de resultados de um agente em um certo cenário. Sutton (1992) revela que o conceito de *reinforcement learning* é um tanto quanto antigo. As primeiras pesquisas publicadas na área foram de Minsky (1961), seguido de Waltz e Fu (1965). Apesar disso, foi apenas no final da década de 80 que estudos relevantes na área voltaram a ser publicados, incluindo Chris Watkins (1989) e Paul Werbos (1987), que linkaram o assunto com programação dinâmica, está intimamente ligado à inteligência artificial e aprendizado de máquina.

Segundo Honda et al. (2017), aprendizado de máquina é um programa que busca melhorar uma performance de um agente, através da realização de uma tarefa e adquirindo experiência. Ainda segundo os autores, existem três tipos de aprendizado de máquina.

O primeiro deles é o aprendizado supervisionado, na qual, através de uma série de variáveis independentes, tentamos prever uma variável dependente. Pode-se pensar, como um exemplo, em uma análise de regressão linear, da qual quanto mais dados tem, melhor será a previsão da variável dependente. Segundo Castro (2017), a principal diferença desse tipo de aprendizado para o aprendizado por reforço é que o primeiro possui os pares de input/output no início do estudo. Dessa forma, ele precisa definir quais ações tomar e em seguida medir qual o impacto delas no resultado.

O segundo é o aprendizado não supervisionado, no qual as variáveis independentes não são tão bem definidas, por diversos motivos. Dessa forma, o algoritmo precisa achar

padrões que permitam que ele crie uma representação informativa dos dados. Um exemplo é o perfil de um cliente de um determinado estabelecimento. A priori não conhece os dados do cliente específico, mas olhando as transações de compras de muitos recibos, pode-se auferir algumas relações que ajudem a tomar decisões. Esse tipo de sistema é difícil de avaliar, pois não tem acesso aos dados iniciais e como eles foram definidos.

Por último está o aprendizado por reforço, que utiliza a tentativa e erro, baseada em uma função de entrada, e avalia os resultados de cada uma das ações tomadas. O processo se dá por meio de atribuição de recompensas e punições para cada decisão, sem necessidade de determinar necessariamente como a tarefa deve ser feita. Desta maneira, este é um método que permite que o algoritmo faça tentativa e erro, e que seja punido quando erra, e premiado quando acerta. É por meio desse tipo de algoritmos que se desenvolveram os robôs que jogam xadrez contra profissionais, por exemplo.

Segundo Dias Júnior (2012), nesse método, o agente está inserido em um ambiente e reage a ele com algumas possíveis ações. No primeiro momento, o agente se encontra em um estado inicial, antes da ação ser tomada. Após a ação, esse estado se altera. Esse novo estado gera um feedback determinado por um valor para o algoritmo, para que ele discerne, segundo um critério pré-definido, se aquele output foi desejável ou não, que se denomina de reforço. O reforço normalmente é dado por $\{0,1\}$ ou então por números reais. A depender do resultado dado pelo reforço, o algoritmo vai definir maior ou menor probabilidade de voltar a tomar aquela ação quando se encontrar novamente naquele estado inicial. Em seguida, o ciclo se repete, para continuar o aprendizado, como visto na Figura 6.

Figura 6– Funcionamento do *reinforcement learning*



Elaborado por Dias Júnior (2012).

O objetivo final do agente é encontrar uma política de ações que gere o melhor resultado no longo prazo. “The agent's goal is to find a policy and state-update function so as

to maximize the expected sum of discounted rewards” (MURPHY, 1998, p.1). Está intrínseco a esse modelo uma abordagem de exploração. Desta forma, caso na tentativa e erro ele encontre um parâmetro que entregue um resultado desejável, ele tenderá a utilizar novamente essa ação. Entretanto, pode ser que haja uma outra ação que entregue um resultado ainda mais favorável, e que não será atingido porque o sistema acredita que já entregou o melhor resultado. Assim, torna-se necessário que há um mecanismo de ações aleatórias de exploração, para que ele realmente busque entre uma grande gama de possibilidades de ações, e encontra o resultado ótimo.

Balancear esse mecanismo é essencial para a qualidade do modelo, pois caso a aleatoriedade seja alta demais, o aprendizado adquirido poderá não ser bem aproveitado. No caso contrário, de ser baixa demais, podemos convergir para um resultado subótimo. É importante ressaltar que o *reinforcement learning* geralmente é aplicado em sistemas não determinísticos, ou seja, que uma ação tomada mais de uma vez poderá entregar resultados diferentes. É o caso, por exemplo, de uma definição de critérios de ativos a comporem um portfólio, pois as características e comportamento daquele ativo podem variar com o tempo.

Reinforcement Learning é uma abordagem ainda em construção, portanto, a quantia de trabalhos com aplicação em economia é pequena. Um exemplo é Du, Zhai e Lv (2016), que em seu estudo realizaram um modelo que escolhia se comprava ou vendia ativos baseado em *momentum*, ou seja, se o ativo havia no estado anterior apresentado movimento positivo ou negativo. O algoritmo aprendia conforme ia tomando decisões em cada um dos ativos a cada momento, utilizando uma técnica denominada *recurrentreinforcement learning*. Os resultados alcançados mostram que a aprendizagem de máquina simplifica a busca por maximização de retorno e evita problemas de maximização dinâmica.

Existem alguns exemplos de *reinforcement learning* aplicados a finanças, como por exemplo, Giannoccaro e Pontrandolfo (2002). Os autores utilizaram a ferramenta para ajudar na tomada de decisões relacionada à cadeia de suprimentos. Para tal, utilizaram o processo de decisão de Markov, que é uma extensão da cadeia de Markov clássica com ações e recompensas. Os autores salientam que o *reinforcement learning*, possibilitaram encontrar uma política de inventário mais próxima do ótimo.

O presente estudo propõe desenvolver um modelo de otimização de alocação portfólio de criptomoedas que se baseia no uso de algoritmos de otimização com restrições e limites (bandas) para os parâmetros a serem alterados pelo algoritmo. Esses limites serão variáveis e controlados por um método de *reinforcement learning*. A cada período de tempo, o agente do modelo consegue se adaptar ao mercado que está em constante mutação.

3.ESCOLHA DA CARTEIRA DE CRIPTOMOEDAS BASEADO NO APRENDIZADO POR REFORÇO

3.1 SELEÇÃO DAS CRIPTOMOEDAS

Para Gil (2008), uma pesquisa experimental determina um objeto de estudo, as variáveis que influenciam ela, define as formas de controle e de observação dos efeitos sobre o objeto estudado. No caso deste trabalho, o objeto de estudo sobre o qual será realizado um experimento é uma carteira fictícia de criptomoedas, com a finalidade de observar o seu retorno em um certo período de tempo.

Considere um agente tomador de preços que busca alocar seu patrimônio, inicialmente em \$1.000.000,00, entre cinco diferentes criptomoedas. O critério de elegibilidade foi sua capitalização. Segundo dados do CoinMarketcap (2018), as cinco criptomoedas de maior capitalização são: *Bitcoin* \$142.478.043.110; *Ethereum* \$53.085.811.446; *Ripple* \$26.970.313.426; *Bitcoin Cash* \$16.502.623.412; *Litecoin*. \$8.735.794.413. Estas criptomoedas citadas, representam mais de 85% da capitalização em circulação (Tabela 1 do anexo apresenta a capitalização das principais moedas).

A cada período de tempo t , o agente aloca seu patrimônio de tal forma que a soma dos módulos dos pesos de cada moeda seja equivalente a 100%. Precisamente, o agente poderá, por meio de utilização de contratos futuros, realizar operações de compra (acredita que o preço irá subir) ou operações de venda (sua previsão indica que o preço irá cair). Para realizar estas operações, deverá deixar o dinheiro equivalente ao total operado de garantia, de forma que o patrimônio não poderá ser alavancado¹.

Além da definição da posição a ser efetuada (comprado/vendido), o agente precisa definir a intensidade de cada negociação, ou seja, escolher a porcentagem do patrimônio que deve ser alocado em cada criptomoeda. Assumindo que o agente tem a intenção de diversificar seu portfólio, estabelece as bandas (limites) máximos e mínimos da porcentagem do patrimônio que pode ser alocada em cada criptomoeda. Como se verá adiante, a aprendizagem por reforço implementada, aumenta (reduz) as bandas caso a estratégia anterior tenha obtido sucesso (fracasso). Por fim, ressalta-se que os custos de transação associados a cada realocação de carteira são incrementados diariamente. Desta forma, ao final de cada dia, é possível calcular o patrimônio líquido total da carteira.

¹ Não permitir que o agente alavanque seu patrimônio, possibilita que os retornos das carteiras de criptomoedas tenham uma maior confiabilidade em sua comparação.

3.2 DADOS, COMPOSIÇÃO DO PORTFÓLIO E O ALGORITMO DE APRENDIZAGEM POR REFORÇO

No presente trabalho tomar-se-á como as séries empíricas observadas, os preços reais em dólares das criptomoedas no período de janeiro a março de 2018, todos dados diários². A escolha do período foi devido à escassez de dados de todas as moedas, visto que algumas não existiam até metade do ano de 2017. Além disso, decidimos eliminar o período final do ano de 2017 visto que foi um período bastante atípico no mercado, devido à viralização do conceito de criptomoedas na mídia mundial. Uma vez obtido o preço diário das criptomoedas, é possível calcular o retorno das criptomoedas. Formalmente:

$$R_t = \sum_{i=1}^5 (P_{i,t} - P_{i,(t-1)}) * Q_{i,t}, \quad (3)$$

no qual R denota o retorno do ativo; P é o preço; Q é a quantidade e t é o período de análise.

Visando deixar o modelo o mais realista possível, adotar-se-á os custos de negociação da *exchange* BitMex, uma das maiores corretoras de criptomoedas do mundo. É importante ressaltar que independente se a operação for de compra ou de venda, o custo de negociação incidirá. O custo de transação (CT) de uma ordem é de 0,05% do valor da ordem. Desta forma, CT é dado por:

$$CT_t = \sum_{i=1}^5 (|Q_{i,t} - Q_{i,(t-1)}| * P_{i,t} * 0,0005), \quad (4)$$

Para todo $t = 0$, ou seja, no período inicial, o patrimônio da carteira (PL) é expresso por:

$$PL_0 = \sum_{i=1}^5 Q_i * P_{i,t}, \quad (5)$$

Por sua vez, o patrimônio líquido da carteira no período ($t \geq 1$) é dado por:

$$PL_t = PL_{t-1} + R_t - CT_t. \quad (6)$$

O processo das revisões das estratégias ocorre entre o fim do período t e o período t+1. Neste período, os agentes avaliam os desempenhos passados de suas estratégias (alocações) e buscam as estratégias que lhe oferecem as melhores recompensas dentro das possíveis bandas. Neste processo adaptativo, os agentes utilizam do mecanismo de aprendizado por reforço para reavaliar e escolher as melhores bandas de alocação.

No processo de aprendizagem por reforço, como destacam ROTH & EREV (1995, p. 165) uma estratégia de decisão que apresentou uma recompensa positiva no período t, têm sua probabilidade de ser escolhida no período t+1 reforçada. Entretanto, caso o resultado do

² Os dados empíricos foram obtidos em CoinMarketCap (2018).

payoff tenha sido indesejável no período t , a probabilidade desta estratégia ser escolhida no período $t+1$ é enfraquecida. Formalmente, esta relação é expressa por:

$$(1+0.10) \text{ se } P_t > P_{(t-1)}, \quad (7)$$

$$(1-0.10) \text{ se } P_t < P_{(t-1)}. \quad (8)$$

Pela função de atualização dos limites de alocação, pode-se inferir que dado uma estratégia que resulte em um retorno positivo no período t , a banda de alocação será ampliada para o período $t+1$. Todavia, se o preço do ativo selecionado reduzir no período t , sua banda de alocação será reduzida no período $t+1$, e desta forma, o tomador de decisão tende a alocar menos recurso no ativo que gerou menor recompensa.

Uma vez definido as novas bandas para o período seguinte, é utilizado o algoritmo de otimização SLSQP para encontrar a melhor composição de pesos de cada moeda na carteira de forma a maximizar o retorno no período presente, e estes pesos serão alocados no processo de seleção dos ativos no período $t+1$.

Precisamente, o mecanismo de otimização adotado busca obter o maior no patrimônio do dia anterior, e retornar quais os pesos ótimos de cada ativo. Esse resultado é obtido através de diversas tentativas de distribuições de pesos e análise de seus resultados, para em seguida retornar a que obteve o melhor. Essas tentativas seguem algumas restrições, a saber: a primeira restrição é com relação à soma dos pesos. A soma do módulo dos pesos dos ativos (comprado/vendido) deverá ser igual a 1 (ou 100%).

A segunda restrição é com relação as bandas (limites) mínimo e máximo de alocação em cada peso. Todas as moedas iniciam o estudo com os limites de (-40%, 40%). Entretanto, esses limites se alteram devido ao mecanismo de aprendizado por reforço implementado. Uma vez definido a alocação ótima para cada ativo no próximo período, inicia-se uma nova etapa.

3.2.1 Implementação Computacional

Para implementar computacionalmente o modelo de escolha do agente, far-se-á uso do software Jupyter Notebook, que é um editor de textos para linguagem de programação Python. Lembra-se que a cada período de tempo t , o agente irá se deparar com o mesmo conjunto de criptomoedas e deverá escolher uma combinação (compra/venda) de tal forma que todo o seu patrimônio seja utilizado.

Como forma de definir as condições iniciais do modelo e evitar possíveis vieses de seleção, para o período $t=0$, assume-se que as cinco criptomoedas são adotadas de maneira

uniforme (20% do patrimônio em cada ativo) e apenas a posição comprada é a possível de ser efetuada neste período inicial.

Após a designação das condições iniciais, é possível obter o retorno do ativo (equação 3) e posteriormente calcular os custos da operação (equação 4). Posteriormente, utiliza-se o algoritmo de otimização SLSQP, para encontrar a melhor composição de pesos de cada moeda na carteira de forma a maximizar o retorno no período presente, e estes pesos serão alocados no processo de seleção dos ativos no próximo período. Assim, torna-se possível calcular o patrimônio líquido da carteira (equação 6).

Todo o procedimento aqui descrito para a escolha dos pesos ótimos dos ativos no período $t=1$, aplica-se para qualquer outro período $t \geq 2$.

4. RESULTADOS

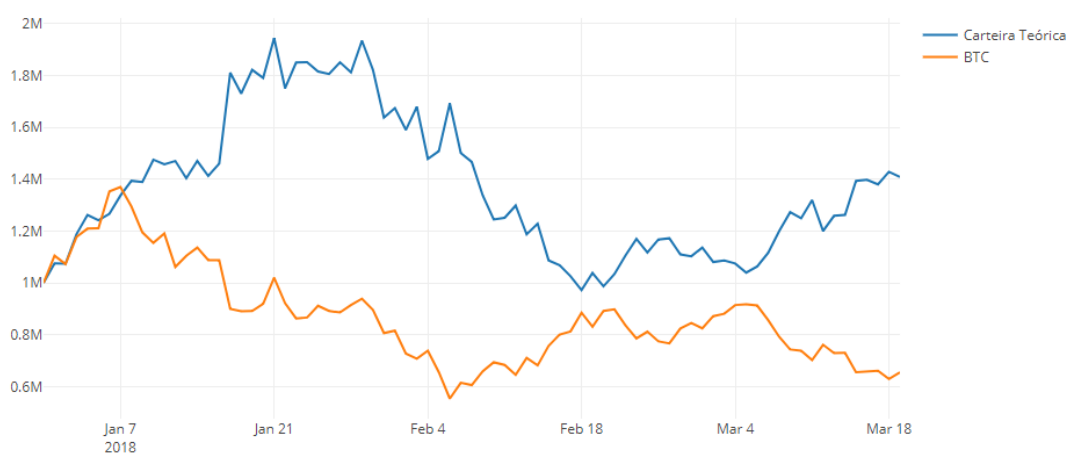
4.1 RESULTADOS DA CARTEIRA

Para aplicação da presente análise foi tomado como base um patrimônio inicial, em 01/01/2018, de R\$ 1.000.000,00. Posteriormente, observa-se como foi o desempenho da carteira teórica desde essa data até 19/03/2018. Foram calculados para o período, tanto para a carteira teórica quanto para o Bitcoin (BTC), os seguintes indicadores: o retorno financeiro acumulado, o retorno percentual acumulado, os custos totais, a variância, o maior *drawdown* diário e o retorno médio diário.

O retorno financeiro foi calculado utilizando os dados de patrimônio líquido das carteiras nas respectivas datas, dado pelas fórmulas 6, sendo assim o retorno financeiro acumulado é o da data final da série. O retorno percentual foi obtido dividindo o valor do patrimônio líquido da respectiva data pelo patrimônio inicial, sendo o retorno percentual acumulado o da data final da série. Os custos totais são a soma de todos os custos de transação ao longo da série (equação 4). A variância é calculada utilizando os dados de retorno diário. O maior *drawdown* diário é o valor mínimo entre os retornos diários.

Na Figura 7, compara-se o patrimônio alocado na carteira teórica (em azul) contra o patrimônio alocado exclusivamente em Bitcoin.

Figura 7– Carteira teórica versus Bitcoin



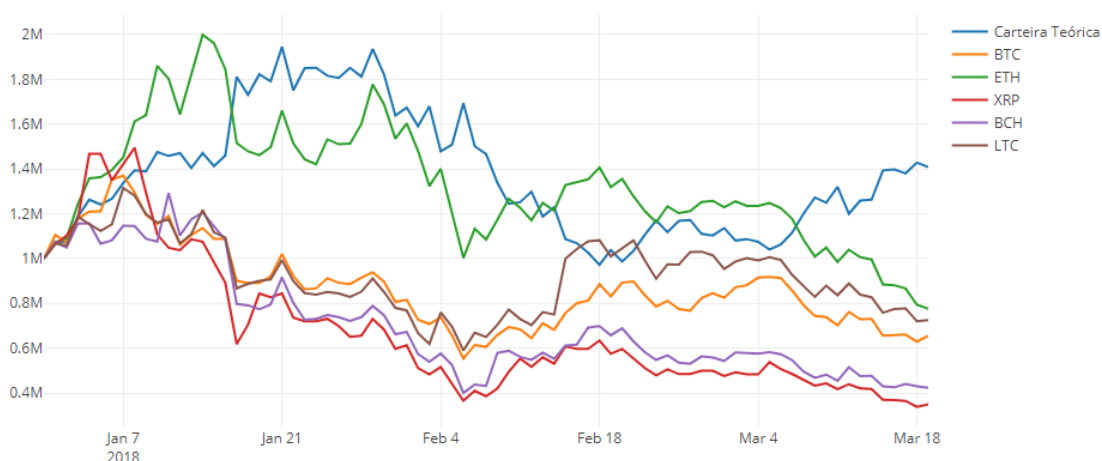
Elaboração própria.

Como pode-se observar, a carteira teórica se mantém constantemente acima do Bitcoin. Ao final do período, o retorno financeiro acumulado da carteira teórica é de um

ganho de R\$ 455.508,57, ao mesmo tempo em que o retorno financeiro acumulado do Bitcoin é de uma perda de R\$ 344.421,05. Isso representa um retorno de 45,55% para a carteira teórica contra -34,44% para o Bitcoin.

É importante ressaltar que esse retorno já está contabilizando os custos envolvidos na realocação das carteiras, que totalizaram R\$ 86.006,12. Visto que se adquirindo apenas Bitcoin com o patrimônio inicial não haveriam realocações, o custo nesse caso é zero. Não é somente com relação ao Bitcoin que a carteira teórica se comportou de forma superior. Analisando o comportamento das outras criptomoedas envolvidas na carteira (Figura 8), pode-se observar que, apesar de ter momentos em que Ethereum (ETH) ter sido superior, ao final do período todas as criptomoedas terminaram com um resultado pior do que a carteira proposta.

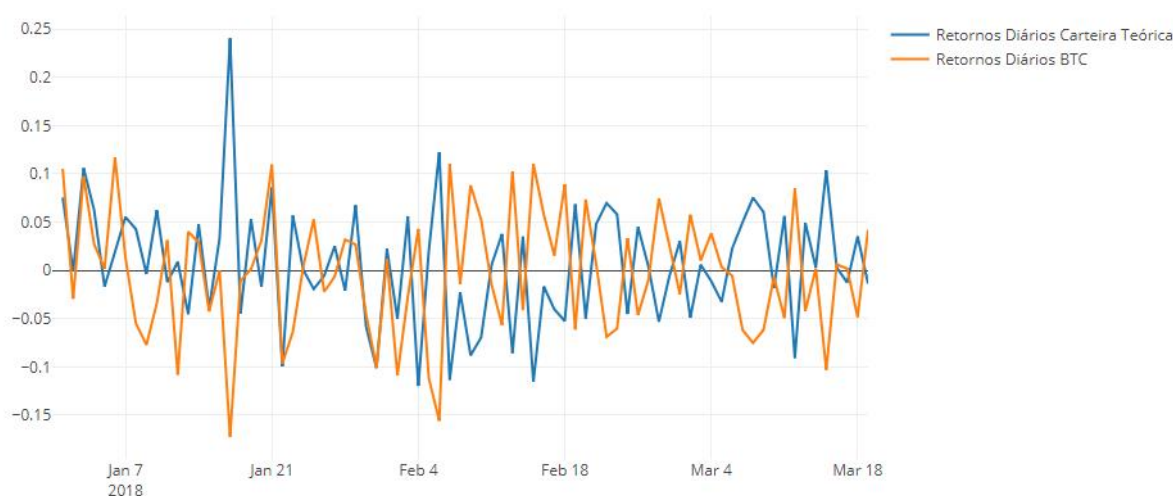
Figura 8– Carteira teórica versus as cinco maiores criptomoedas



Elaboração própria.

Com relação às variáveis associadas ao risco, analisamos a variância e o *drawdown* máximo da carteira contra o Bitcoin. No período estudado, a variância da carteira proposta foi de 0,59% e a do Bitcoin foi de 0,41%, demonstrando que a carteira teve um risco maior. Apesar disso, o *drawdown*(maior queda em um dia) máximo diário foi maior no caso do Bitcoin do que da carteira teórica. Enquanto a carteira chegou a perder 15,74% do seu patrimônio em um dia, o Bitcoin chegou a perder 17,23% em um único dia. Como medida de análise dia a dia dos retornos dos ativos, utilizou-se o retorno médio diário, como visto na Figura 9.

Figura 9– Retornos diários da carteira teórica versus Bitcoin



Elaboração própria.

O resultado médio do período, para a carteira, foi bastante superior ao do Bitcoin, ficando em 0,76% para o primeiro contra -0,33% para o segundo. Pode-se observar na Figura 9 que os retornos entre os dois ativos eram em certa proporção inversamente correlacionados na maior parte dos dias. Pelos resultados obtidos, se conclui que para o período estudado o retorno de um valor aplicado na carteira teórica teria sido superior ao de qualquer outra criptomoeda.

Fazendo uma análise comparativa com o Bitcoin nesse mesmo período verifica-se duas coisas. Primeiro, que o máximo *drawdown* da carteira teórica se demonstrou inferior ao do Bitcoin. Isso demonstra que a carteira teve nesse período um prejuízo diário máximo menor que o do Bitcoin. Segundo lugar, que ao ponto em que o retorno do Bitcoin foi negativo, a carteira não só obteve retorno superior ao ativo, mas foi positivo em 45,55%, mesmo envolvendo custos de transação.

4.2 COMPARAÇÕES COM OUTRAS ESTRATÉGIAS

Foram realizados, para efeito comparativo, testes com mais duas estratégias. Assim, estes testes têm por objetivo avaliar como os recursos utilizados para tomada de decisão de alocação impactaram nos resultados obtidos.

A primeira delas é mantendo os limites de alocação fixos em -40% até 40%, ou seja, sem o mecanismo de *reinforcement learning*. O resultado dessa estratégia em comparação com o Bitcoin pode ser visto na Figura 10.

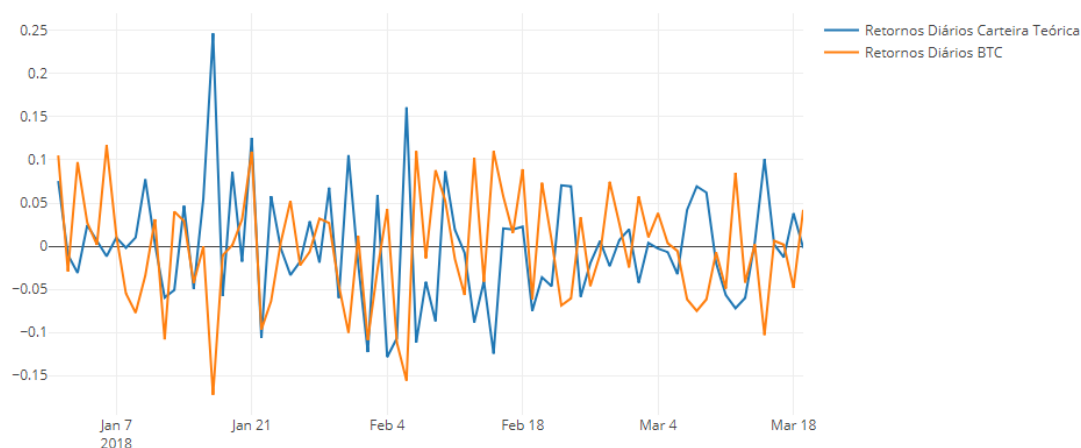
Figura 10– Resultado com bandas de 40%

Elaboração própria.

Esta estratégia teve retornos inferiores, quando comparado com *reinforcement learning* (Figura 7), demonstrando o valor do *reinforcement learning* para melhorar as decisões de alocações com o passar do tempo. O retorno acumulado do período foi de uma perda de R\$ 206.582,75, mais próximo do resultado do Bitcoin apresentado na subseção anterior. Esse resultado equivale a uma rentabilidade percentual de -20,66%.

Apesar da variância nesse caso ter sido inferior à carteira teórica com *reinforcement learning*, segue acima do Bitcoin, estando agora em 0,43%. Também houve uma diminuição pouco significativa no *drawdown* diário máximo, agora em 12,84%. A figura 11 mostra os retornos para cada dia da série para o modelo com bandas de 40% para a carteira e o Bitcoin. O retorno médio diário foi negativo, de cerca de -0,08%, contra -0,33% do BTC.

Figura 11– Retornos com bandas de 40%



Elaboração própria.

O segundo modelo comparativo é uma carteira com pesos fixos (20% para cada moeda). Esse é o modelo de alocação mais ingênuo entre os testados, visto que além de não aprender com as decisões corretas e erradas anteriores, não toma decisões de alocação. Na Figura 12 pode-se observar o comportamento dessa estratégia contra o Bitcoin.

Figura 12– Pesos fixos para cada criptomoeda

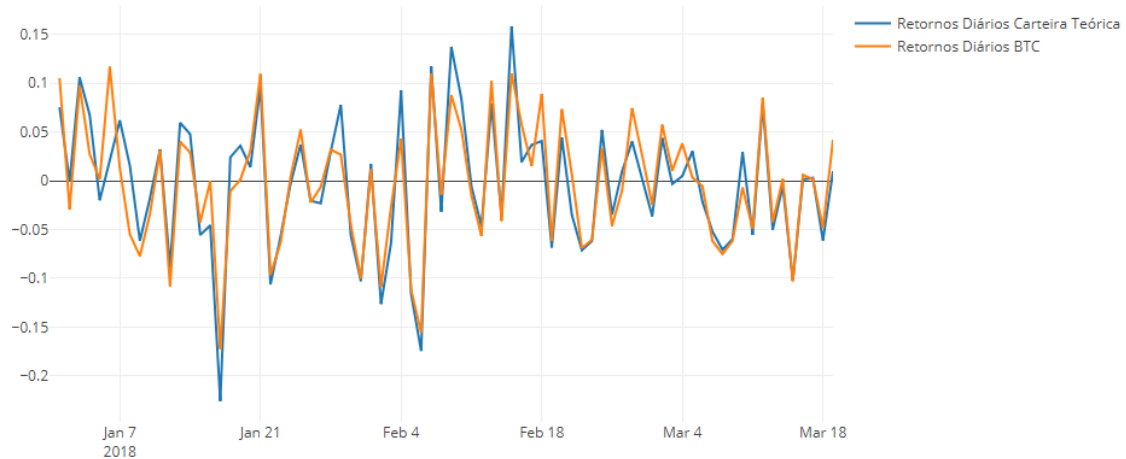


Elaboração própria.

O retorno financeiro acumulado do período é de uma perda de US\$ 410.371,29, superior inclusive à perda do Bitcoin no mesmo período. Ainda assim, obteve uma variância maior que esse ativo, de cerca de 0,46%. A maior perda diária da carteira foi de 22,59% do patrimônio, bastante superior ao do Bitcoin e a maior entre as estratégias analisadas. Na Figura 13 pode-se ver os retornos para cada dia da série para a estratégia com pesos fixos em

20% e para o BTC.

Figura 13– Retornos diários com pesos fixos



Elaboração própria.

O retorno médio dessa estratégia foi de -0,43%, sendo esse o pior entre todas as estratégias. O máximo *drawdown* da carteira foi maior do que o do modelo proposto. Uma observação interessante que se pode extrair do gráfico de distribuição dos retornos diários entre a carteira de o Bitcoin é que, com pesos fixos em 20% por moeda, os retornos diários são muito próximos dos do Bitcoin, demonstrando alta correlação do restante das moedas com ele.

Para uma análise mais abrangente da comparação entre as estratégias mais ingênuas e a carteira teórica, será utilizada uma metodologia própria inspirada no índice de Sharpe, bastante conhecido no mercado financeiro tradicional. Precisamente, e dado por:

$$IR = \frac{RC - RB}{VC}, \quad (9)$$

no qual IR é o Índice de Risco/Retorno; RC o Retorno da Carteira; RB o Retorno do Bitcoin; e VC a Variância da Carteira.

Para a carteira teórica com *reinforcement learning*, na qual a rentabilidade no período é de 45,55%, rentabilidade do Bitcoin de -34,44% e variância da carteira de 0,59%, temos:

$$IR = \frac{0,4555 - (-0,3444)}{0,0059}$$

$$IR = 135,58$$

Para a carteira teórica comparativa sem *reinforcement learning*, na qual a rentabilidade no período é de -20,66%, rentabilidade do Bitcoin de -34,44% e variância da carteira de 0,43%, temos:

$$IR = \frac{(-0,2066) - (-0,3444)}{0,0043}$$

$$IR = 32,05$$

Para a carteira teórica comparativa com pesos fixos, na qual a rentabilidade no período é de -41,03%, rentabilidade do Bitcoin de -34,44% e variância da carteira de 0,46%, temos:

$$IR = \frac{(-0,4103) - (-0,3444)}{0,0046}$$

$$IR = -14,33$$

Dessa forma, observa-se que analisando o risco retorno por essa metodologia, a carteira proposta foi muito superior às carteiras comparativas, ademais, o modelo que utiliza pesos fixos obteve risco retorno negativo. Isto ocorre porque, apesar do estudo não focar na redução do risco, e sim no aumento do retorno, houve pouca variação na variância da carteira proposta em relação ao Bitcoin, mas um grande aumento no retorno esperado.

A Tabela 2 sintetiza os indicadores para cada carteira do capítulo anterior.

Tabela 2 - Indicadores por carteira

Carteira	Maior <i>drawdown</i>	Retorno	Variância	IR
<i>Machine learning</i>	15,74%	45,55%	0,59%	135,58
Bandas de 40%	12,84%	-20,66%	0,43%	32,05
Pesos fixos de 20%	22,59%	-41,04%	0,46%	-14,33
Apenas Bitcoin	17,23%	-34,44%	0,41%	-

Elaboração própria.

5. CONSIDERAÇÕES FINAIS

Nos últimos anos, o advento das criptomoedas, como o Bitcoin, chamou a atenção dos *players* do mercado financeiro. Criptomoedas são ativos financeiros que podem dar alto retorno, e vale a pena explorar o tema a partir de uma abordagem de otimização de portfólio, a partir das ideias de Markowitz (1952). A carteira de criptomoedas foi composta pelas cinco maiores criptomoedas em capitalização: Bitcoin, Ethereum, Ripple, Bitcoin Cash e Litecoin.

Para encontrar o resultado ótimo de uma carteira de criptomoedas, o método escolhido foi um modelo baseado em agente com *reinforcement learning*, que recompensa acertos e pune erros na direção do mercado. A carteira fictícia começou com um patrimônio líquido de US\$ 1.000.000,00, com o estudo ocorrendo de 01/01/2018 até 19/03/2018.

Ao final do período do estudo, o patrimônio líquido da carteira foi maior do que teria sido se todo o valor fosse investido em uma das cinco criptomoedas individuais. A única que em alguns períodos teve desempenho melhor foi a Ethereum, mas no final ela acabou perdendo para a carteira teórica assim como outras. Quanto ao risco, a carteira apresentou ligeiramente mais risco do que o Bitcoin, mas mesmo assim o maior *drawdown* diário foi do Bitcoin.

Também foram realizadas comparações com outras estratégias. Primeiro, uma carteira sem *reinforcement learning*, com bandas de 40%. Para esta estratégia, o resultado foi inferior ao da carteira teórica, embora ainda melhor do que investir todo o patrimônio líquido em Bitcoin. Outro modelo comparado foi uma alocação com pesos ingênuos – fixos em 20% para cada moeda. Esta foi a pior das alocações em todas as medidas. Um fato interessante é que os retornos e a variação desta estratégia foram muito semelhantes ao do Bitcoin, mostrando alta correlação entre as criptomoedas.

Para análise do risco e retorno, o trabalho utilizou um índice próprio com metodologia semelhante ao famoso índice de Sharpe, aqui denotado como IR.

Assim, observou-se que a carteira inteligente que utiliza *reinforcement learning* possui um índice de risco retorno muito melhor do que as outras carteiras (o Bitcoin foi usado como retorno base para o índice, portanto não apresentando seu próprio IR). A carteira com os pesos mais ingênuos inclusive apresenta IR negativo. Apesar do trabalho focar na maximização do retorno, observa-se com base nos resultados disponíveis que não existe uma grande diferença de risco entre as opções, enquanto o retorno aumenta bastante.

Por fim, vale reiterar que o trabalho demonstra a utilidade da utilização de técnicas computacionais como o *machine learning* na resolução de problemas econômicos, principalmente na área de finanças, permitindo ao investidor explorar novos territórios muito

melhor do que investidores mais ingênuos. O mercado de capitais é análogo a outros ambientes de risco, como o pôquer, onde é muito importante sempre estar à frente da curva. E o uso de um agente sob *reinforcement learning* melhora o seu retorno em relação ao restante do mercado, assim sendo excelente opção para modelar a alocação de recursos em criptomoedas.

REFERÊNCIAS

- ATHEY, Susan. **The Impact of Machine Learning on Economics**. NBER Working Papers, jan 2018.
- BOYD, Stephen; BUSSETI, Enzo; DIAMOND, Steven; KAHN, R.N. **Multi-Period Trading via Convex Optimization**. Foundations and Trends in Optimization, v. 3, n. 1, p. 1-76, 2017.
- CASTELLANO, Murilo. **A Teoria do Risco Financeiro – Parte II**. Disponível em: <<http://www.blogdocastellano.com.br/?p=1643>>. Acesso em: 01 jun 2018.
- CASTRO, U.C. **Derivação de modelos de trading de alta frequência em juros utilizando aprendizado por reforço**, 2017. Dissertação (Mestrado em Economia e Finanças) – Mestrado Profissional em Finanças e Economia, Fundação Getúlio Vargas, São Paulo.
- CHIU, Jonathan; KOEPPL, T.V. **The Economics of Cryptocurrencies – Bitcoin and Beyond**, 2017. Disponível em: <https://www.chapman.edu/research/institutes-and-centers/economic-science-institute/_files/ifree-papers-and-photos/koeppel-april2017.pdf>. Acesso em: 01 jun 2018.
- CHOI, J.J.; LAIBSON, David; MADRIAN, B.C.; METRICK, Andrew. **Reinforcement Learning and Savings Behavior**. The Journal of Finance, v. 64, n. 6, p. 2515-2534, dez 2009.
- CoinMarketCap**: Cryptocurrency Market Capitalizations. Disponível em: <<https://coinmarketcap.com>>. Acesso em: 01 jun 2018.
- DIAS JÚNIOR, E.P.F. **Aprendizado por reforço sobre o problema de revisitação de páginas web**, 2012. Dissertação (Mestrado em Informática) – PPG em Informática, Pontifícia Universidade Católica, Rio de Janeiro.
- DU, X.; ZHAI, J.; LV, K. **Algorithm trading using q-learning and recurrent reinforcement learning**. Citeseer, v. 1, p. 1, 2016.
- GIANNOCARO, Ilaria; PONTRANDOLFO, Pierpaolo. **Inventory management in supply chains: a reinforcement learning approach**. International Journal of Production Economics v. 78, n. 2, p. 153-161, jul 2002.
- GIL, A. C. **Métodos e Técnicas de Pesquisa Social**. 6ª. ed. São Paulo: Editora Atlas, 2008.
- HONDA, H., FACURE, M., PENG, Y. **The three types of machine learning**, 2017. Disponível em: <<https://lamfo-unb.github.io/2017/07/27/tres-tipos-em-english>>. Acesso em: 01 jun 2018.
- KITSCHKE, J.H. **Investigation of the Convergence Behaviour of Different Optimization Algorithms using Shape Optimization in Flow Problems**, 2014. Disponível em: <http://www.mathcces.rwth-aachen.de/_media/3teaching/0classes/archiv/037_ces_seminar_convergenceoptimizationalgorithms.pdf>. Acesso em: 01 jun 2018.

KRINK, Thiemo; PATERLINI, Sandra. **Differential Evolution for Multiobjective Portfolio Optimization**. RECent Working Papers, jun 2008.

MARKOWITZ, Harry. **Portfolio Selection**. Journal of Finance, v. 7, p. 77- 91, mar 1952.

MINSKY, M.L. **Steps toward artificial intelligence**. Proceedings I.R.E., v. 49, p. 8-30, 1961.

MULLAINAHAN, Sendhil; SPIESS, Jann. **Machine Learning: An Applied Econometric Approach**. Journal of Economic Perspectives, v. 31 n. 2, p. 87-106, mar-jun 2017.

MURPHY, **A brief introduction to reinforcement learning**, 1998. Disponível em: <<https://www.cs.ubc.ca/~murphyk/Bayes/pomdp.html>>. Acesso em: 01 jun 2018.

NAKAMOTO, Satoshi. **Bitcoin: A Peer-to-Peer Electronic Cash System**, 2009. Disponível em: <<https://bitcoin.org/bitcoin.pdf>>. Acesso em: 01 jun 2018.

NARAYANAN, Arvind; BONNEAU, Joseph; FELTEN, Edward; MILLER, Andrew; GOLDFEDER, Steven. **Bitcoin and Cryptocurrency Technologies: a comprehensive introduction**. Princeton University Press, 2016. 336 p.

PEREIRA, L.B.T.; HENRIQUE, D.C. **Otimização de investimentos pelo modelo de Markowitz via desenvolvimento de uma ferramenta em Excel**. Iberoamerican Journal of Industrial Engineering, v. 8, n. 16, 2016.

ROTH, A. E.; EREV, I. **Learning in extensive form games: Experimental data and simple dynamic models in the intermediate run**. Games and Economic Behavior, v. 8, p. 164-212, 1995.

SANTOS, A.A.P. **Otimização de Carteiras baseada em Modelos de Correlações Condicionais**. Análise Econômica, v. 34, n. 65, mar 2016.

SANTOS, J.O. **Análise da relação risco e retorno em carteiras compostas por índices de bolsa de valores de países desenvolvidos e de países emergentes integrantes do bloco econômico BRIC**. Revista Contabilidade & Finanças, v. 21, n. 54, p. 23-37, set-dez 2010.

SUTTON, R.S. **Introduction: The Challenge of Reinforcement Learning**. In: Sutton R.S. **Reinforcement Learning**. The Springer International Series in Engineering and Computer Science (Knowledge Representation, Learning and Expert Systems), v. 173. Boston, MA:Springer, 1992.

WALTZ, M.D., FU, K.S. **A heuristic approach to reinforcement learning control systems**. IEEE Transactions on Automatic Control, v. 10, n. 4, p. 390–398, 1965.

WATKINS, C.J.C.H. **Learning with delayed rewards**, 1989. Tese (Doutorado em Psicologia), Departamento de Psicologia, Cambridge University.

WERBOS, P.J. **Building and understanding adaptive systems: A statistical/numerical approach to factory automation and brain research**. IEEE Transactions on Systems, Man and Cybernetics, v. 17, n. 1, p. 7-20, jan-fev 1987.

YERMACK, David. **Is Bitcoin a Real Currency? An economic appraisal**. NBER Working Papers, dez 2013.

APÊNDICE

Tabela 1: capitalização das principais criptomoedas em 19/03/2018

Nome	Capitalização
<i>Bitcoin</i>	\$142.478.043.110
<i>Ethereum</i>	\$53.085.811.446
<i>Ripple</i>	\$26.970.313.426
<i>Bitcoin Cash</i>	\$16.502.623.412
<i>Litecoin</i>	\$8.735.794.413
Cardano	\$4.623.185.583
NEO	\$4.370.535.000
Stellar	\$4.346.580.385
EOS	\$4.051.168.807
IOTA	\$3.598.324.314
Monero	\$3.335.956.016
Dash	\$3.123.413.425
NEM	\$2.580.021.000
Tether	\$2.221.730.295
TRON	\$2.050.469.154
VeChain	\$1.894.484.288
Ethereum Classic	\$1.774.030.874
Lisk	\$1.280.295.749
Qtum	\$1.236.419.465
OmiseGO	\$1.089.467.508

Elaboração própria baseada em dados de CoinMarketCap (2018).