A Proposal For Risk Distribution Over Intelligent Credit Networks

Gaurav Singhal* Dr Laurence E. Day†
{gaurav, laurence}@arboreum.dev

Draft Version 0.1.0

March 9, 2020

Abstract

Peer to peer lending is a rapidly growing market and is expected to reach US$800 billion by 2025. However, peer-to-peer lending is currently fraught with risk: risk assessment itself is tricky, loans often have no collateral, and the mechanisms for enforcement are often weak.

In a nutshell, we are creating an upgraded peer to peer system that significantly derisks loans using artificial intelligence and decentralized networks, where the amount of a loan held by any one lender is repeatedly novated amongst colleagues who have established explicit lines of trust. Repeated across a network, any one loan becomes 'fragmented', reducing the exposure that any one person bears, with local AI agents performing optimal allocations for each person based on their risk appetite and dollar value in play to maximise their risk adjusted return.

In this brief overview, we will explain the underlying theory and concepts behind this model, and illustrate the computational difficulties which require further attention in order to be viable for real-life application.

* *MSc Information Science & Economics, University of Michigan*
† *PhD Computer Science, The University of Nottingham*

Contents

1 Network Behaviour

Our goal is to create a variant of the peer-to-peer lending model which we refer to as *network lending*. In this model, agents who are connected to each other via a network represented as a weighted directed acyclic graph (Thyfronitis Litos & Zindros, 2017) distribute funds in an approximately Pareto efficient manner (Mock, 2011) to agents demanding said funds (borrowers) through other agents on the network (lenders).

1.1 Network Characteristics

The basic characteristics of these agents and the network are:

- Agents (representing the vertices V of the network graph) first deposit funds (Q_*) that are held in a global escrow account.
- Any two agents $A, B \in V$ can be connected to each other by a directed weighted edge $e_{AB} \in E$ where the weight $w_{e_{AB}}$ denotes the credit capacity offered by A to B.
- The credit capacity of an edge must be less than or equal to the funds deposited by the agent at the tail of the edge: i.e. $w_{e_{A\cdot}} \leq Q_A \; \forall \; e_{A\cdot}$.
- We define 'trust' to be a synonym for the (extension of) credit capacity between two distinct agents.
- The existence of an edge does not imply the existence of a transpose: i.e. $e_{AB} \not\Rightarrow \exists \; e_{BA}$.
- The credit capacities of an edge and its transpose need not be equal: i.e. $w_{e_{AB}} \neq w_{e_{BA}}$.
- Since trust (as defined above) is not reciprocal, an agent A has two sets of neighbours:
  1. Trustors ($\{ \; B \parallel \text{B} \in V, \; e_{BA} \in E\}$): the set of vertices that explicitly trust agent A, providing incoming lines of credit.
  2. Trustees ($\{ \; B \parallel \text{B} \in V, \; e_{AB} \in E\}$): the set of vertices that agent A explicitly trusts, and to whom flow outgoing lines of credit.
  3. The trustor/trustee relationship is an inverse one - if agent B is a trustor of agent A, then agent A must be a *trustee* of agent B, and vice versa.

1.2 Agent Characteristics

Any given agent A has a known *risk of default* (p_A) on any funds which they have borrowed, represented as a beta distribution parameterised by scale terms α_A and β_A. Only members of an agents trustor set are able to directly view these parameters; other agents can only infer them by asking other agents until the request reaches a trustor. The resulting message is then propagated back to the requestor, not unlike in a game of telephone.

Hence, any agent B that is not a trustor of A has a *subjective view* of the default risk posed by A, described by parameters α_{BA} and β_{BA} (Sun, Yu, Han, & Liu, 2006).

Further characteristics that are local to agents are:

- Each agent A has a known risk tolerance represented by two parameters:
  1. z_A - the maximum allowable loss on deposited funds Q_A, e.g. 1%.
  2. k_A - the maximum probability that projected losses exceed z_A, e.g. 5% probability that losses exceed 1%.
- A third parameter λ_A representing the expectation-variance trade-off can be inferred from the above two parameters.
- Borrowers must pay some interest rate r on their loan (parameterised by the loan amount D and the repayment period T), and they can offer a percentage of the loan amount c as collateral.
- Each agent wishes to maximise the expected value of their profits from lending their funds to other agents on the network whilst satisfying the above risk parameters.
- Each agent is limited to interacting solely with their trustor/trustee sets. No agent has a global view of the network, and hence all funds, messages and responses must be routed through neighbours.

The (subjective) risk for a given lender increases as their distance from a borrower increases. Further, no funds can be routed at a loss. Ergo, incentive compatibility requires a loan to be either derisked or yield higher interest as it is propagated from primary lenders (direct trustors of the borrower) to secondary and higher lenders (trustors of primary lenders et al). This lends itself to our first lemma:

Lemma 1.1. *All payment paths from lenders to a given borrower must satisfy risk monotonicity.*

Loan propagation occurs through *novation*, which we discuss in the next section.

2 Agent Behaviour

2.1 Novation

Novation is the central behaviour of agents on a network such that we have defined above; it is the process by which funds are propagated from borrowers to many lenders on said network. In a financial context, novation is the act of purchasing debt, amending the contract and then reselling some or all of the debt to additional parties, as illustrated below:

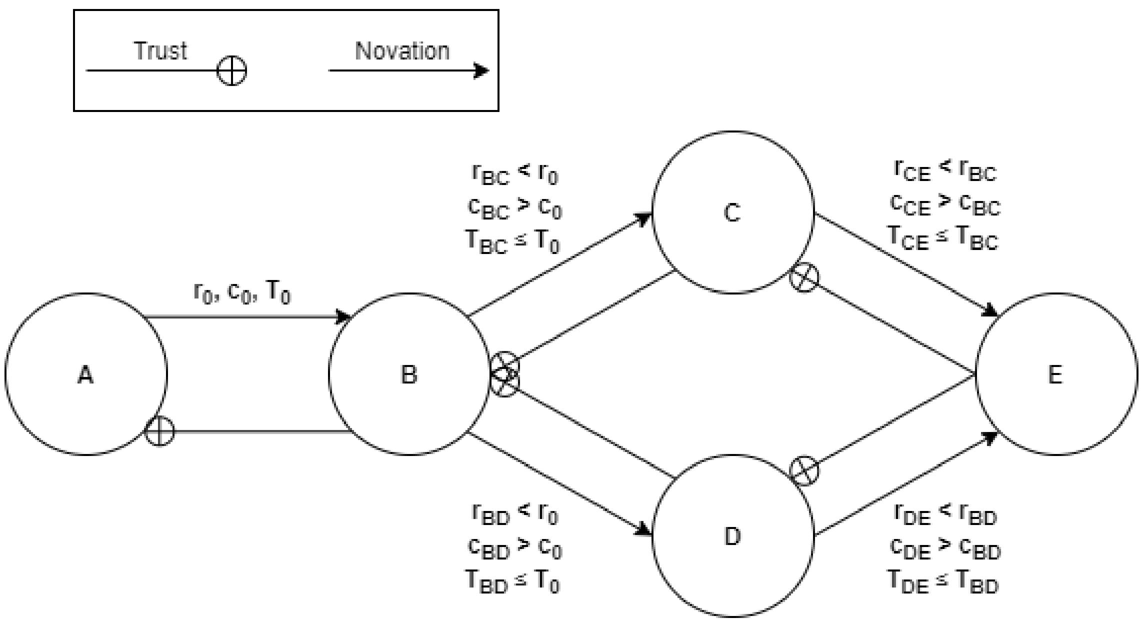

Figure 1: The Novation Process

In Figure 1 above, agent B purchases debt from a borrower agent A in their capacity as a trustor of A. As B is a trustee for agents C and D, they sell part of this debt to both of these trustor agents (provided the terms are within their risk profile) at a decreased interest rate and collateral, for a time period less than or equal to than the initial loan. In turn, agents C and D are both trustees of agent E, and so sell parts of their own newly-purchased debt to E on their own terms (a further decreased interest rate etc.). At settlement of the initial loan from B to A, provided that no default occurs, all secondary agents involved in the chain profit from the difference in terms.

With respect to the aforementioned credit network, lenders always perceive borrowers as riskier - due to the dispersion of their perceived risk profile - the further they are from the borrower on the network, per Lemma 1.1. As rational novation cannot occur at a loss, an additional lemma emerges from incentive compatibility:

Lemma 2.1. *For a loan originator to make a profit, $r_n \leq r_{n+1}$. As interest cannot increase, Lemma* 1.1 *necessitates that the collateral on the novated debt monotonically increases and/or the tenor decreases as we move further away from the borrower, i.e. $c_n > c_{n+1}$ and $T_n \geq T_{n+1}$.*

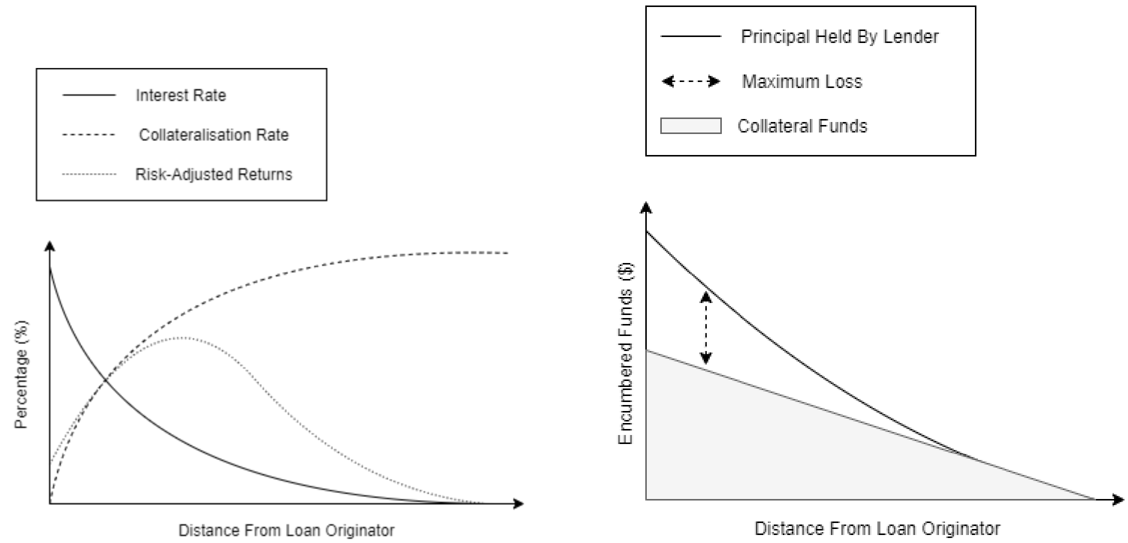

Figure 2: Consequences Of Lemma 2.1 As Functions Of Distance From Loan Originator

In Figure 1 we saw a relatively simple example of a loan being propagated by novation. Consider agent E, the agent at the end of a novation chain. This agent has no trustors that they can novate debt further to, and so from their perspective, the issue of participating in the original loan becomes a tractable optimisation problem, which we describe in Section 4.

3 An Extended Example

The underlying concepts explained thus far are best illustrated by seeing a loan - and its novation - in action. Consider the network below:

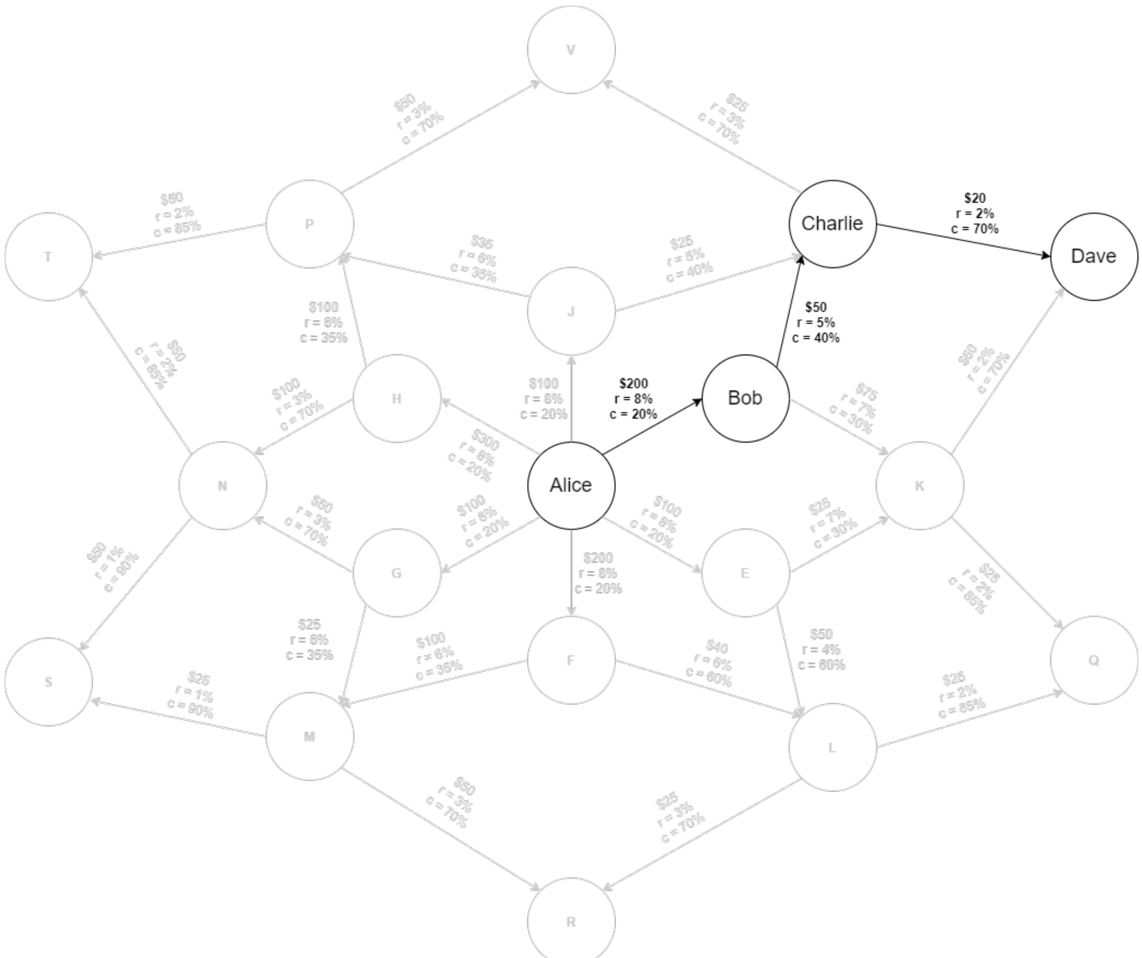

Figure 3: Explicit Novation Chain From Borrower Alice To Lender Dave

In this network, a borrower Alice has requested (and received) a loan for $1,000. Her loan was initially purchased by six of her trustors, one of whom is Bob. Bob is in turn trusted by Charlie, who is trusted by Dave. Whilst Figure 3 above shows example paths for all of Alice's participating trustors, we will be explicitly focusing on the chain novation from Alice through to Dave, exploring how each participant profits from their involvement.

All agents on the above network have already performed the prerequisite steps for novation; namely, they have:

1. Assessed the risk that Alice poses to them, building a *subjective default risk profile*.
2. Coalesced their trustors' demand functions (if any).
3. Broadcast their own demand functions (for Alice's loan) to their trustees.

3.1 Following Alice Down The Rabbit Hole

Based on the results that Alice received from her trustors, she requested that her loan of \$1,000 is taken with an interest rate of 8% and a 20% collateral deposit, and of this principal, Bob contributed \$200. Now, we describe the process by which we propagate this segment of the loan through Bob's own network of trustors.

Bob - who has knowledge of the demand for Alice's loan from Charlie, sells him \$50 of this newly acquired \$200 debt. This \$50 is sold at a lower interest rate than the initial debt (5% rather than 8%), but Bob has doubled the collateralisation rate, bringing it up to 40%. Here is where Bob and Charlie now stand overall:

- Bob holds \$150 of Alice's debt, at 8% interest with 20% collateralisation.
- Charlie holds \$50 of Alice's debt, at 5% interest with 40% collateralisation.
- Bob now also holds a *swap note* on Alice's debt worth \$50 at 3% interest (8% - 5%) provided Alice repays her loan in full, but may cost him up to \$10 (\$50 × (40% - 20%)) if Alice defaults.

Let us consider what this change in circumstance has done to Bob's potential profits:

- Bob initially placed \$160 at risk (\$200 × (100% - 20%)) in order to earn \$16 in interest, yielding an *effective* interest rate (EIR) of \$16/\$160 = 10%.
- After novating \$50 of this \$200 debt to Charlie, Bob is now risking \$130 (the result of \$150 × (100% - 20%) + \$50 × (40% - 20%)) in order to earn \$13.50 in interest (\$150 × 8% + \$50 × (8% - 5%)), yielding an EIR of \$13.50/\$130 = 10.38%.

It seems counterintuitive, but this novation has *increased* Bob's overall risk-adjusted return, whilst simultaneously exposing Charlie to an asset at terms he is comfortable with.

We now repeat this process with Charlie, who – knowing his trustor Dave's demand for Alice's principal – novates \$20 of the \$50 debt he has purchased from Bob at a further reduced interest rate of 2%, adding an additional 30% collateral.

As was the case with Bob, we compare Charlie's pre-and-post-novation returns, noting that Charlie now holds a swap note worth \$30 at 3% interest with a maximum potential loss of \$9 (\$30 × (70% - 40%)):

- Pre-novation: Charlie risks \$30 (\$50 × (100% - 40%)) to earn \$2.50 in interest, for an effective interest rate of 8.33%.
- Post-novation: Charlie now risks \$24 (\$30 × (100% - 40%) + \$20 × (70% - 40%)) in order to earn \$2.10 in interest (\$30 × 5% + \$20 × (5% - 2%)), yielding an EIR of \$2.10/\$24 = 8.75%.

As was the case with Bob, Charlie has reduced his exposure whilst increasing his risk-adjusted returns. Finally, Dave - the end of the chain - now holds an asset representing a debt of \$20 in order to earn \$0.40 in interest, with a maximum potential loss of \$6.

As a rule, novation always increases risk-adjusted return, allowing loan risk to propagate throughout a network as a consequence of agents making rational, self-interested decisions.

3.2 The Bigger Picture

In the interests of completeness, below we provide a table of the *true* - considering all chains - figures of funds exposed, interest and returns pre-and-post-novation for *all* loan participants, leaving the intermediate calculations to the interested reader:

Participant	Pre-Novation Funds Exposed	Post-Novation Funds Exposed	Interest	Risk-Adj. Return
Bob	\$160.00	\$77.50	\$8.25	10.64%
E	\$80.00	\$42.50	\$4.25	10.00%
F	\$160.00	\$79.00	\$7.60	9.62%
G	\$80.00	\$48.75	\$5.00	10.25%
H	\$240.00	\$145.00	\$15.00	10.34%
J	\$80.00	\$42.25	\$4.65	11.00%
Charlie	-	\$31.50	\$2.60	8.25%
K	-	\$51.25	\$5.50	10.73%
L	-	\$24.75	\$3.15	12.72%
M	-	\$63.75	\$5.75	9.01%
N	-	\$32.50	\$3.00	9.23%
P	-	\$65.25	\$5.60	8.58%
Dave	-	\$21.00	\$1.40	6.67%
Q	-	\$7.50	\$1.00	13.34%
R	-	\$22.50	\$2.25	10.00%
S	-	\$7.50	\$0.75	10.00%
T	-	\$15.00	\$2.00	13.34%
V	-	\$22.50	\$2.25	10.00%
Total	**\$800.00**	**\$800.00**	**\$80.00**	-

As we can see, the funds exposed - loan principal not covered by collateral - and interest due do not change, but are rather dissipated amongst the various network participants. Moreover, the novation of Alice's loan permits twelve network participants to act as lenders despite not knowing Alice directly.

4 Portfolio Optimisation

We now return our focus to the issue of network participants deciding how much debt they are willing to purchase when presented with a new lending opportunity. The setup for this problem is as follows:

- A lending agent A has a maximum amount of funds q_{AB} that they are willing to allocate to the debt of a borrower agent B, where $q_{AB} \leq Q_A$.
- A has a current portfolio Γ consisting of multiple risky assets $\gamma_{1,\ldots,N}$. Each constituent asset γ_i is parameterised by the following:
  - x_i - the amount of the asset held in the portfolio.
  - r_i - the interest rate of the asset.
  - c_i - percentage of collateral covering the asset in the event of default.
  - T_i - tenor of asset (time remaining until funds are returned).
  - α_i, β_i - scaling terms for the beta distribution parameterising risk of default.
- A has known risk aversion parameters z_A, k_A and λ_A.

The lender must decide how much of a borrower agent B's debt they wish to add to their portfolio as a new asset γ_{N+1}, where $\gamma_{N+1} = x_{AB} \leq q_{AB}$, given r_{AB}, c_{AB}, T_{AB}, α_{AB} and β_{AB}.

4.1 Simplifying Assumptions

With the following two assumptions, we can approach the above problem - the calculation of γ_{N+1} - in a similar manner to traditional Markowitz portfolio optimisation (Markowitz, 1952).

Assumption 1: without loss of generality, we can approximate the *beta distribution* via the *Kumaraswamy distribution* (Jones, 2009), which is defined as -

$$K(x) = ab\ x^{a-1}(1 - x^a)^{b-1}$$

- where $a > 0$, $b > 0$ are the scale parameters of the distribution. Approximating the beta distribution in this way circumvents the fact that its analytical presentation has no closed form and is thus undifferentiable.

Assumption 2: we ignore time dynamics and assume that all debt has the same settlement time (i.e. $T_* = 1$), thus eliminating tenor from the equations to follow.

4.2 The Problem Made Concrete

In making the two assumptions from the previous subsection, we arrive at the following optimisation problem that a lender with no trustors in a novation chain (i.e. Dave in Section 3) must solve.

An agent A seeking to optimise their portfolio aims to maximise -

$\Gamma = \texttt{max}(\theta^\intercal(PR + (1-P)C))$, such that:

(1) $b \times B(1+\frac{1}{a}, b) = \frac{\theta^\intercal(PR+(1-P)C)}{\theta^\intercal R}$

(2) $b \times (B(1+\frac{2}{a}, b) - B(1+\frac{1}{a}, b)) = (\frac{\Sigma(R-C)^\intercal\theta}{\theta^\intercal 1_M})^\intercal \times \Omega \times \frac{\Sigma(R-C)^\intercal\theta}{\theta^\intercal 1_M}$

(3) $(1-(1-k_A)^{\frac{1}{b}})^{\frac{1}{a}} \geq z_A \times \frac{\theta^\intercal}{\theta^\intercal R}$

- where

- θ is the vector $[x_1, \ldots, x_M]$ representing the total amount of each risky asset γ_i in Γ.
- P is the vector $[1 - p_i, \ldots, 1 - p_M]$ representing the payback probability (the default complement) of each risky asset γ_i in Γ.
- R is the vector $[1 + r_1, \ldots, 1 + r_M]$ representing the total percentage return realised if a given risky asset γ_i is repaid in full.
- C is the vector $[c_1, \ldots, c_M]$ representing the percentage of collateral in place for each risky asset γ_i in the event of default.
- a and b are the parameters of the Kumaraswamy distribution.
- $B(\cdot, \cdot)$ is the Euler integration *beta function* - not to be confused with the beta distribution mentioned in Assumption 1 of the previous subsection - defined as:

$$B(x, y) = \int_0^1 t^{x-1}(1-t)^{y-1}\, dt$$

- Σ is the vector $[\sigma_1, \ldots, \sigma_M]$ of standard deviations of each risky asset γ_i, where:

$$\sigma_i = \sqrt{\alpha_i\beta_i} \times \frac{1}{\alpha_i+\beta_i} \times \sqrt{1+\alpha_i+\beta_i}$$

- 1_M is the vector containing M iterations of the value 1.
- Ω is the $M \times M$ correlation matrix of risky assets in Γ.
- z_A and k_A are the risk aversion parameters of agent A.

The objective is simply to maximise the expected value of the portfolio Γ. In the event that no defaults occur and all risky assets are repaid in full, the returns on Γ are equal to the dot product $\theta \cdot R$. In contrast, if *every* constituent asset defaults, the 'returns' are simply the aggregate of all collateral, $\theta \cdot C$.

Equations (1) and (2) seek to establish equality between the beta distributions and their corresponding Kumaraswamy distributions. The left hand sides of these equations correlate to the first and second moments of the Kumaraswamy distribution, and the right hand sides to the empirical normalised first and second moments of the portfolio Γ itself. Equation (3) imposes limits on the allowable tail risk of the resulting Kumaraswamy distribution describing the portfolio Γ.

Assuming that a portfolio Γ is optimised for an existing M risky assets, the problem of deciding how much debt agent A should purchase from a potential borrowing agent B becomes one of solving for the unknown θ vector element $x_{M+1} = x_{AB}$ as well as the associated vector entries - i.e. c_{M+1}, r_{M+1} - which ensure that Γ *remains* optimal.

To do this, we seek to frame the above in terms of a *demand function* D_{AB} such that -

$$x_{AB} = D_{AB}(r_{AB}, c_{AB})$$

- where $r_{AB} > 0$ and $0 < c_{AB} < 1$.

5 The Demand Function & Message Passing

The ability of agents to pass these demand functions as messages to their neighbours is at the heart of the propagation-by-novation process. If a potential novating agent B receives a demand function D_{CA} from a trustor C in the form of a message, the problem becomes one of calculating terms $r_{BA}^{Out} = r_{CA}^{In}$ and c_{CA}^{In} at which agent B would be able to novate some of the *initial borrowing agent* A's incoming debt x_0 to C, where $\{r, c\}_{BA}^{Out} = \{r, c\}_{CA}^{In}$. We can illustrate this as follows:

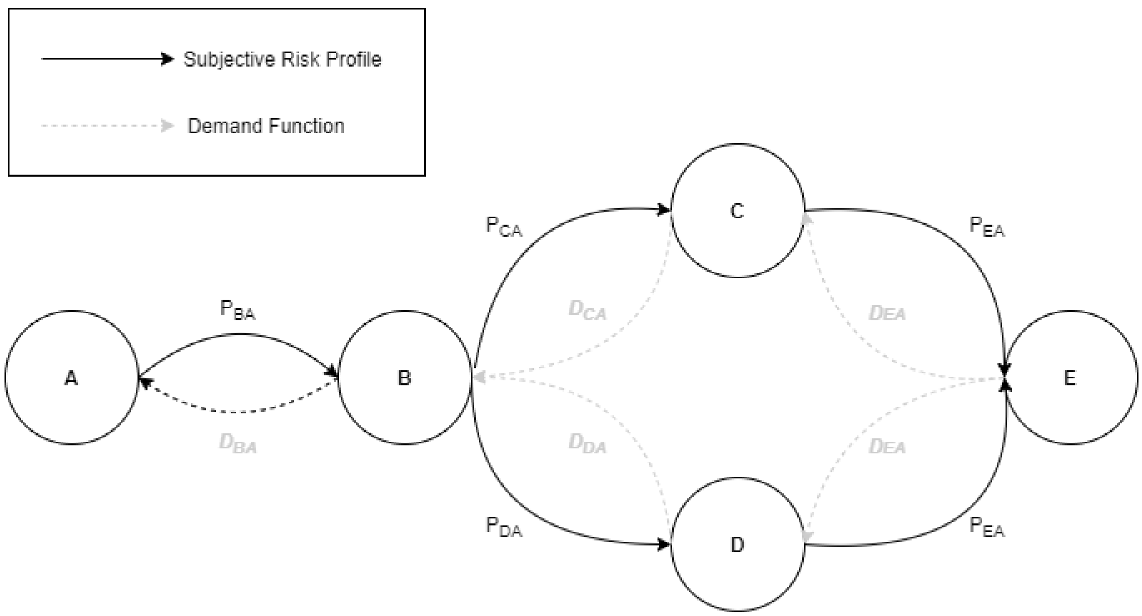

Figure 4: Demand Function Back-Propagation

In response to receiving a demand function from a trustor (i.e. D_{CA}), agent B solves for three unknowns x_{BA}, r_{BA}^{Out} and c_{BA}^{Out} and extends their portfolio vectors as follows:

- $P = [\dots,\ 1 - p_{BA},\ 1 - p_{BA}]$ - agent B is effectively creating two new risky assets in their portfolio; the first the debt which B purchases from A, and the second the debt which B sells at novated terms to C. p_{BA} is the probability of default that B assesses A to possess.

- $R = [\dots,\ 1 + r_{BA}^{In},\ 1 + r_{BA}^{In} - r_{BA}^{Out}]$ - since agent B is selling debt to a willing agent C, the first entry corresponds to the interest rate offered by A, and the second indicates that the difference between the buy and sell rates is the expected rate of return they will earn on the novation.

- $C = [\dots,\ c_{BA}^{In},\ -c_{BA}^{Out}]$ - the first entry corresponds to the collateral of the purchased debt, whereas the second is negative as its encumbered by the additional collateral they have added during the novation process.

- $\theta = [\ldots, x_{BA}, D_{CA}(r^{In}_{CA}, c^{In}_{BA} + c^{In}_{CA})]$ - where D_{CA} is the demand function of C for the debt of borrower A and $r^{In}_{CA} = r^{Out}_{BA}$ and $c^{In}_{CA} = c^{Out}_{BA}$).
- $\Sigma = [\ldots, \sigma_{BA}, \sigma_{BA}]$ - the two assets have the same underlying normalized standard deviation.
- $\Omega = \begin{bmatrix} \ddots & \ldots & \ldots & \ldots \\ \vdots & \ddots & \rho_{AB} & -\rho_{BA} \\ \vdots & \rho_{BA} & \ddots & \vdots \\ \vdots & -\rho_{BA} & \ldots & 1 \end{bmatrix}$ - where ρ_{BA} is the portfolio correlation vector.

5.1 Propagation-By-Novation

In light of everything we have encountered so far, we can now outline the full propagation-by-novation process as follows:

Wave 1: a borrowing agent A broadcasts a message requesting a risk assessment for a particular amount of debt. The message propagates away from A through various levels of trustors, where each intermediate agent combines their information about the trustor which delivered the message with agent A's initial message to create a subjective risk profile of A. As the message moves further away from A, uncertainty around the risk increases.

Wave 2: after a finite number of edge traversals, risk assessment stops (when the message cannot propagate to any additional trustors). At this point, agents begin to propagate their demand functions regarding the initial loan amount back towards agent A. It is important that a strict post-order depth-first search ordering is followed to ensure that an agent has all of their neighbours demand functions prior to computing their own.

Wave 3: finally, agent A coalesces the demand functions of their trustors into a function that proscribes the interest rate at which they can borrow the requested amount as a function of collateral. Once A settles on their terms, their trustors - armed with the knowledge of their own trustors demand functions - begin the novation process, propagating the loan and settling accounts. Funds are encumbered when held for collateral, and other funds are channeled to the borrower for the purpose of their loan.

6 Aid Required

Whilst the authors have created a proof of concept (available here) that demonstrates that the underlying concept is sound, there are a number of areas in which further rigour is required, particularly regarding the underlying mathematics and computational optimisation thereof. In order of importance, the following topics must be addressed to move to the next stage of development:

1. Incorporating time (debt tenor) as an element of the portfolio optimisation equations detailed in section 4.2.

2. Incorporating a 'patience' term within the optimisation, factoring in the opportunity cost of lending funds at the first chance, rather than waiting for superior opportunities.

3. A more robust formalisation of the optimisation problem itself: ideally in terms of a modified Weibull distribution (Weibull, 1951) in order to describe the distribution of portfolio returns inclusive of time.

4. At present we solve for demand functions by layering a grid over the curve parameterised by r and c, solving for each coordinate, and applying LOESS (locally estimated scatterplot smoothing) (Cleveland, 1981) over the resulting collection. In practice, the demand function should be both continuous and monotonic: i.e. phrasable via a polynomial equation with a maximum of three terms in any dimension. We welcome expertise on potential ways in which to directly solve for such a polynomial.

5. We are open to alternative ways to solve the problem at hand. There is a significant body of literature on credit networks, but in almost all cases interest rates and collateral are treated as exogenous: they do not arise as a consequence of an agents' perceived risk. Further, credit networks as typically studied are more concerned with contagion, as collateral is usually proffered in the form of other risky assets. In our formulation we do not allow for this, eliminating contagion risk at the expense of capital efficiency.

7 Forthcoming Sections & Update Log

This paper is - in its current form - incomplete. Over time, we will be adding new sections to introduce additional concepts that have thus far been alluded to, or that bring additional clarity. With each update, we will adjust the paper version, and reflect the change in an update log below.

7.1 Forthcoming Sections

- Financial & Trust Impacts Of Default
- Subjective Risk Assessment Via Eigentrust & Loan History
- Fundamentals: Ethereum Message-Passing Framework

7.2 Update Log

Version 0.1.0: version logging and extended example section added.

References

Cleveland, W. S. (1981). LOWESS: A Program for Smoothing Scatterplots by Robust Locally Weighted Regression. *The American Statistician*, *35*(1), 54–54. Retrieved from `http://www.jstor.org/stable/2683591`

Jones, M. (2009, 01). Kumaraswamy's Distribution: A Beta-Type Distribution With Some Tractability Advantages. *Statistical Methodology*, *6*, 70-81. doi: 10.1016/j.stamet.2008.04.001

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, *7*(1), 77–91. Retrieved from `http://www.jstor.org/stable/2975974`

Mock, W. B. T. (2011). Pareto Optimality. In D. K. Chatterjee (Ed.), *Encyclopedia of Global Justice* (pp. 808–809). Dordrecht: Springer Netherlands. Retrieved from `https://doi.org/10.1007/978-1-4020-9160-5_341` doi: 10.1007/978-1-4020-9160-5_341

Sun, Y., Yu, W., Han, Z., & Liu, K. J. R. (2006, 02). Information Theoretic Framework Of Trust Modeling And Evaluation For Ad Hoc Networks. *IEEE Journal on Selected Areas in Communications*, *24*, 305-317. doi: 10.1109/JSAC.2005.861389

Thyfronitis Litos, O. S., & Zindros, D. (2017, Jan). Trust Is Risk: A Decentralized Financial Trust Platform. In (p. 340-356). doi: 10.1007/978-3-319-70972-7_19

Weibull, W. (1951, Jan). A Statistical Distribution Function Of Wide Applicability. *Journal of Applied Mechanics*, *18*, 293-297.