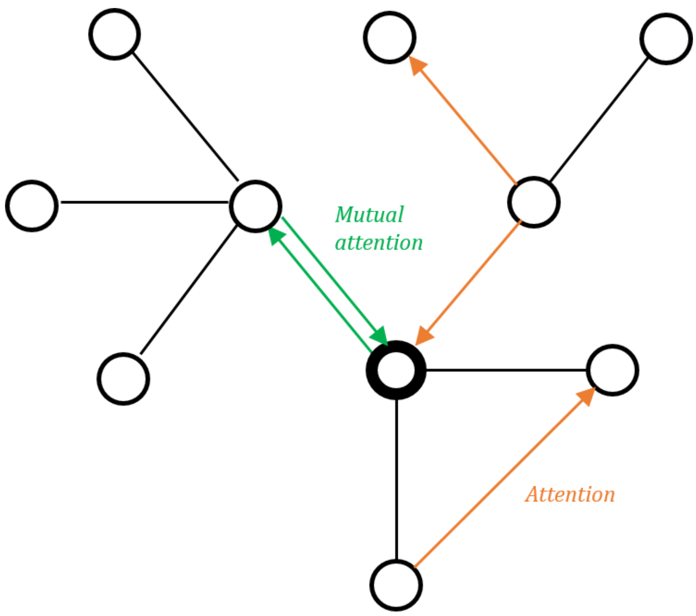
Intima - Mapping Virtual Attention in Networks

Introduction

Communication between people using mobile devices is global and ubiqituous. Almost every person is now at the center of a growing *online network* connecting him/her to contacts using a variety of technologies such as email, phone calls, instant messaging, social media etc. These networks are designed to facilitate a flow of information with certain access mechanisms (e.g. spam filters, add/delete contacts, "friend"-ing etc.)

In this project, we focus on the flow of *virtual attention* along network edges and in particular the degree of *mutual attention* between people/nodes. We observe that in face-to-face communication , mutual attention/gaze is an enabler (and indicator) of effective communication. We are motivated to retrofit/augment online systems with analogous attention signalling mechanisms.

We first describe a model of attention flow using mutual gaze; based on that model, we propose an algorithm/system design for measuring mutual virtual attention in existing online networks (the intima system). We then present some results of simulations and a proof-of-concept; lastly, we introduce some real-world applications of intima.



Intima measures/enables mutual attention in networks

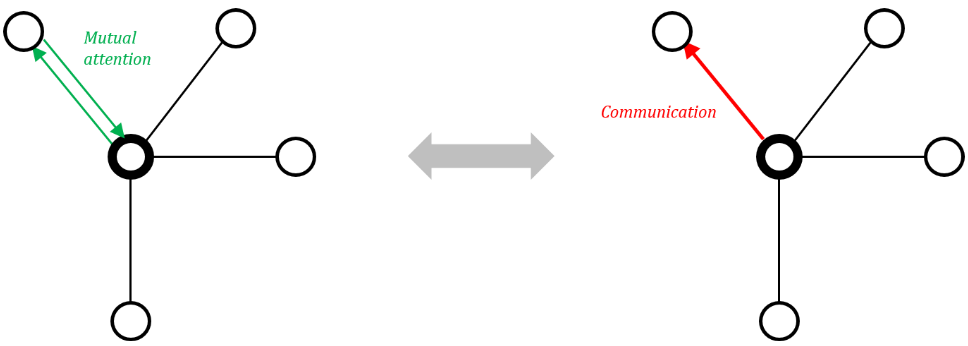
Background

Mutual Attention

In face-to-face communication, attention is a *meta-signal* that *regulates* communication. Attention in the real-world is primarily signalled by the use of *gaze* - we look at each other as we speak. Cognitively, this performs the essential function of focussing on a target and filtering out other signals in noisy environments.

Mutual attention is established by *reciprocation*. *Mutual gaze* in face-to-face communication is a series of "gazes" passed back and forth between two people. This is required as a meta-signal for effective verbal communication - imagine what happens to a conversation if one person abruptly looks away.

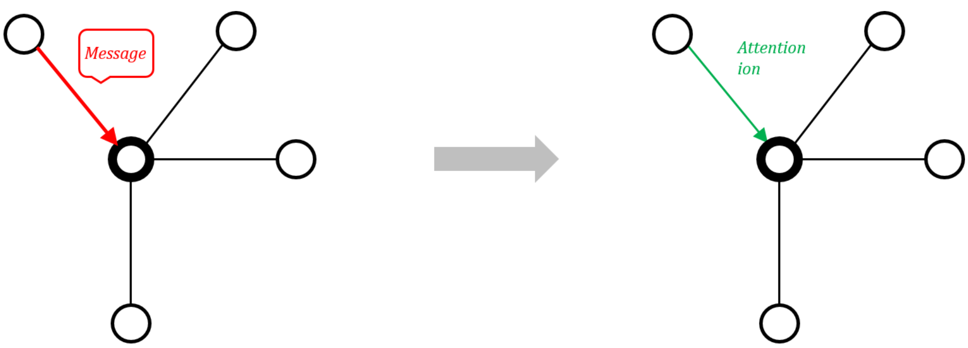
Note that mutual attention and communication are in a feedback loop - each reinforces the other. Mutual attention is needed to initiate communication, and the content of the communication then reinforces and "locks in" mutual attention. Therefore, one can choose to characterize mutual attention as either/both a *measure* and/or an *enabler* of communication.



Feedback between mutual attention and communication

Online Networks

Unlike the face-to-face situation, online communication lacks the non-verbal signalling mechanisms of gaze. However, the sending of a online message from A to B implies *virtual attention*; we might say that attention signalling is *embedded* into communication.



Message sending over online networks implies virtual attention

This allows us to measure the flow of virtual attention by examining the *meta-data* that flows around the communication network - source, destination and timestamp of each message/packet. This extracts the attention signals without having to look at the content of any communication.

Attention Model

In this section, we develop an attention model based on *gaze* in face-to-face communication. This model will inform the implementation of the intima system.

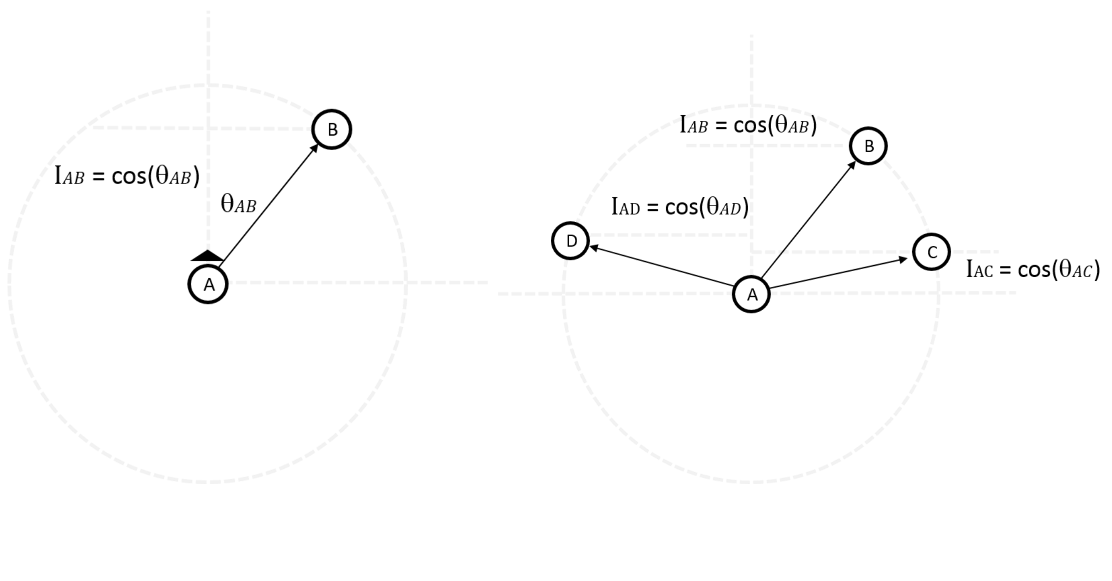
Intensity of Attention

Consider a top-down view of a person A in 2D space. We model a person's centre of gaze as a vector pointing straight ahead with a length of 1. Rotating this vector defines a circle of gaze of radius 1. Now consider a second person B situated on the circle of gaze at angle `theta`. We call this person a *gaze target*.

We introduce the concept of *gaze intensity* (denoted by `I`) which corresponds to the intensity of attention being directed from a person to another. With respect to the reference person, this new person intercepts a gaze intensity of `I` which varies as cosine of the angle of incidence `theta`.

For multiple gaze targets, we can define the total gaze intensity that is emitted by person A (and intercepted by all gaze targets collectively) as simply the sum of intensities:

$$I_A = \sum I_{Ak}$$



Simple geometrical model of gaze in 2D

Conservation of Attention

We now introduce a constraint which we call **conservation of attention**:

At any time, the sum of all attention intensities from one node (to all targets) is less than or equal to a constant maximum value

Or in symbolic terms:

$$I_A(t) < I_{max}$$

The reasoning for this constraint is as follows: attention is a finite resource – we have a certain cognitive capacity for attention which cannot be exceeded. In the presence of multiple targets of gaze, we divide our full capacity of attention between the targets.

At this stage, under the assumption that persons have roughly the same attentional capacity we assume that `I_max` is a constant. Combining with earlier results, this means that for a given person:

$$\sum cos\theta_k \leq I_{max}$$

This cannot be true for arbitrary values of `theta_k`, so to enforce this we must introduce a normalizing factor `K`:

$$I = K \cdot \cos\theta_k$$

and we set the value of `K` such that:

$$if \sum cos\theta_k \leq 1 \ then \ K = 1$$

$$if \sum cos\theta_k > 1 \ then \ K = \frac{I_{max}}{\sum cos\theta_k}$$

So that:

$$K \sum cos\theta_k \leq I_{max}$$

Mutual Attention

For two people gazing at each other, we define the mutual gaze `m` as the pair of intensities:

$$\mathbf{m} = [I_{AB}, I_{BA}]$$

We now introduce two measures derived from the mutual gaze. First, the *skew* `S`:

$$S_{AB} = S_{BA} = \sqrt{|I_{AB}^2 - I_{BA}^2|}$$

Skew S is a measure of the **asymmetry of mutual attention**. Person A may be absorbing a high intensity of gaze from B, but the B may be absorbing very little or zero from A. In the real world, this corresponds to a situation where B is gazing at A, but A is not gazing back.

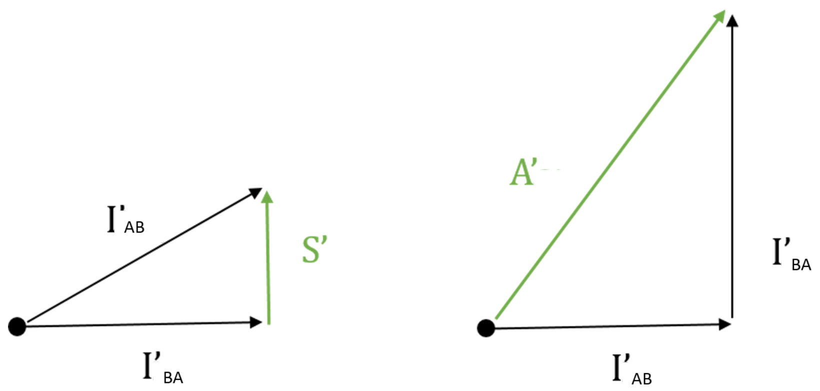
Note that *skew is the inverse of reciprocity*. If intensities in both directions are equal i.e. `I_AB = I_BA` then `S_AB = S_BA = 0` and we say that the mutual gaze is *symmetrical* or *reciprocal*.

Secondly, the *affordance* `A`:

$$A_{AB} = A_{BA} = \sqrt{\frac{I_{AB}^2 + I_{BA}^2}{2}}$$

Affordance A is a measure of the **total intensity of both directions of mutual attention**. It can also be named as the mutual intensity.

The geometrical interpretation of skew and affordance is easy to understand as the opposite side and hypotenuse respectively of triangles where the other two sides are the intensities. [TODO: note the scaling factor here for `A` and amend diagram]



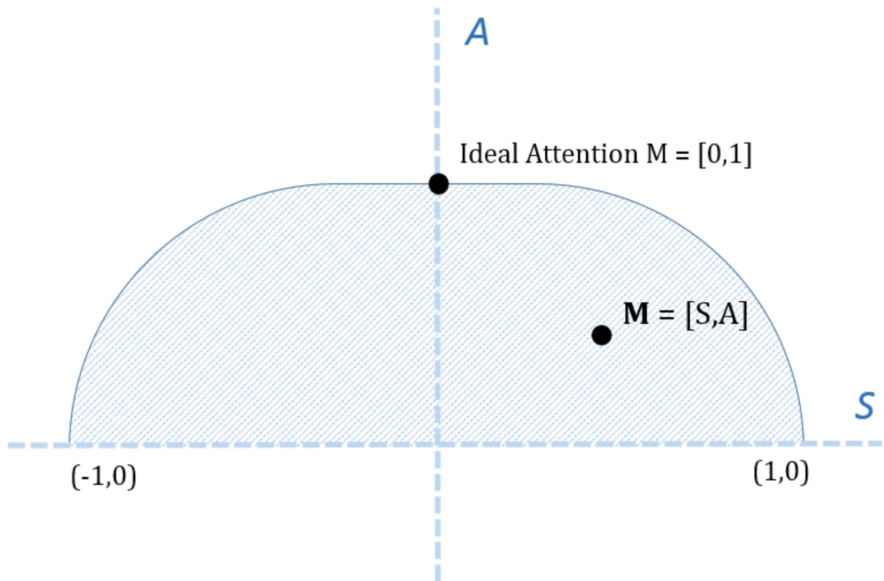
Geometrical interpretation of skew and affordance

Intima in Mutual Attention Space

We now define *intima* `M` as the the pair of affordance and skew:

$$\mathbf{M} = [S, A]$$

Where we constraint `-1 < S < 1` and `0 < A < 1` [TODO: justify]. Intima is a two dimensional number that describes the *quality* of mutual attention. We can plot the value `M` in *mutual attention space* as below.



Mutual attention as a mathematical space

We say that mutual attention is *ideal* when affordance `A` is at a maximum value of 1 (maximum possible mutual gaze intensity) and the skew `S = 0` (complete reciprocity). This corresponds to two persons emitting and absorbing the maximum possible attention in complete symmetry. .

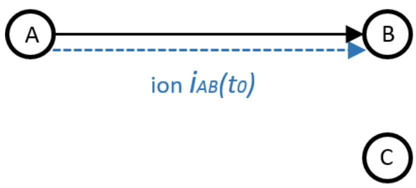
Mutual Virtual Attention in Online Networks

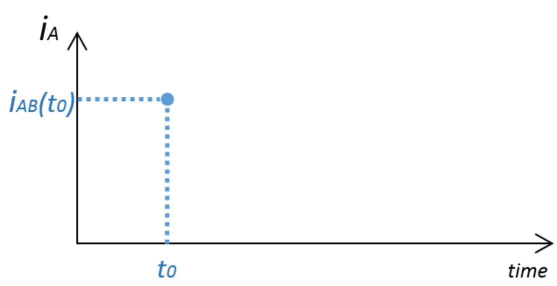
In this section, we present an algorithm and system design that will allow us to a) calculate mutual virtual attention for any two nodes on a network at any point in time given a history of messages and b) send/receive virtual attention signals independently of content messaging. It is based on the attention model in the previous section.

However, the model above is *continuous* and *real-time*. To create a real online system, we need an architecture that is *discrete* and *distributed in time*.

Attention Ions

Suppose that there are three persons on a network, denoted A, B and C, and that A sends an email to B. The act of sending the email implies that person A is directing (virtual) attention towards B at that moment of time. This packet of attention is named an attention *ion*.





Person A sends an attention ion to person B

The ion `i_AB(t)` denotes an attention packet sent from person A to person B at time *t*, with the following data structure:

$$i_{AB}(t) = [id, A, B, i, p]$$

where

`id` is a unique ion id generated at the source
`A, B` are node IDs identifying the source and target of attention
`i` is the attention intensity
`p` is a collection of tags and meta-data

The intensity value `i` expresses the quantum of the attention signal. For example, a single line email may have a smaller intensity than a long email, and a chat message may have a smaller intensity than an email. Ions could even be fired off for every single keystroke when writing an email, but they would then have very small values of `i`.

For shorthand, we use the notation `i_AB(t)` to refer to both the ion itself and the intensity of an ion. Without loss of generality, we restrict `0 < i_AB(t) < 1`.

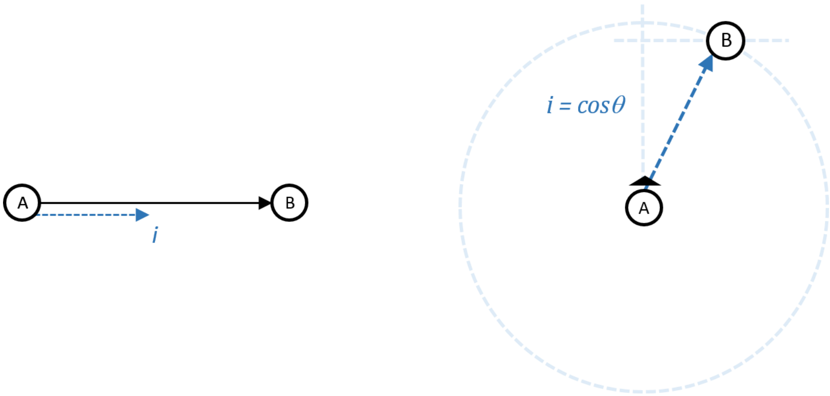
Attention Ions as Gaze

At this point we have a) a model of mutual attention/gaze and b) a method of attention signalling using ions. To synthesize these two models, we make a series of key assumptions. The first of these is the following:

attention ion intensity i is equivalent to gaze intensity i = cos(theta)

Here we are applying a metaphor - sending an attention ion (usually concurrent with a communication) to a target person is mapped to the act of directing gaze towards that person. Comparing to the gaze model, we can see that another way of mapping this quantity is to think of it as `i = cos(theta)` where `theta` is the angle of gaze.

For convenience, we will now use the notation `i` to refer to both attention ion intensity and gaze intensity as interchangeable concepts, and as shorthand refer to 'attention intensity'.



Attention ion intensity equivalent to gaze intensity = 'attention intensity'

Ion Function

Each ion is equivalent to an impulse/delta function in attention intensity, but for a usable model synthesis we need a continuous attention intensity as a function of time. To construct this signal from the discrete attention ions, we need to create a *distribution of attention intensity* in time around the delta function.

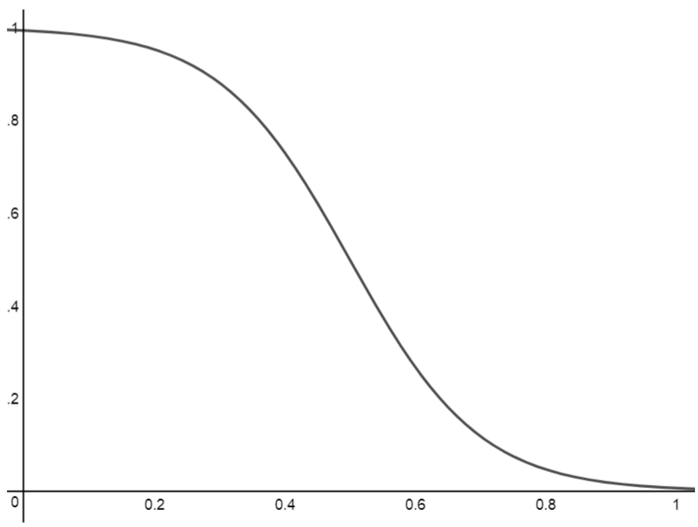
Here we introduce the second of our key assumptions:

the distribution of attention intensity is a descending sigmoid function in time, with an asymptote to zero

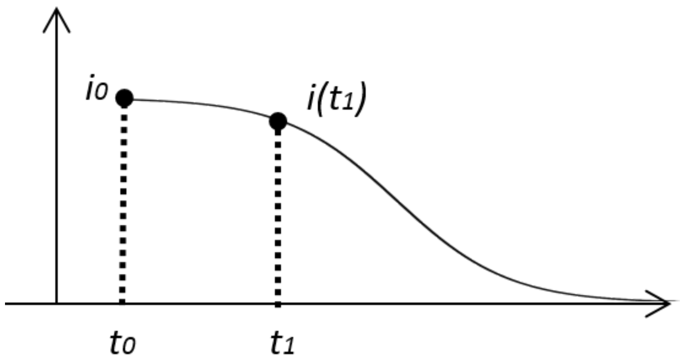
If an ion is emitted at time `t_0` with intensity `i_0`, the distributed attention intensity at a later time `t_1` is:

$$i(t_1) = i_0 \cdot \frac{1}{1 + e^{10\lambda t - 5}}$$

where `lambda` is a decay parameter. For short we call this the ***ion function***.



Ion function shape (i0 = 1 and lambda = 1)

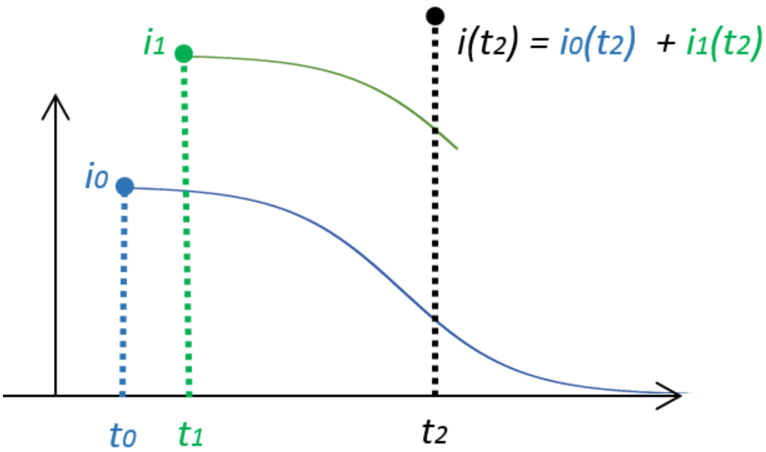


Attention intensity distributed in time from initial ion impulse

We now generalize to the case of multiple attention ions sent over time from the same node. In this case, we introduce another key assumption:

total attention intensity at any time is the sum of ion functions from all previous ions

This gives us a way to calculate total attention intensity emitted by a node at any time as long as we have a record of *all* previous attention ions sent by that node. (In practice, we remove ions from the calculation after a certain time period as their contribution decays to almost zero.)



Attention intensity as a sum of distributions from previous ions

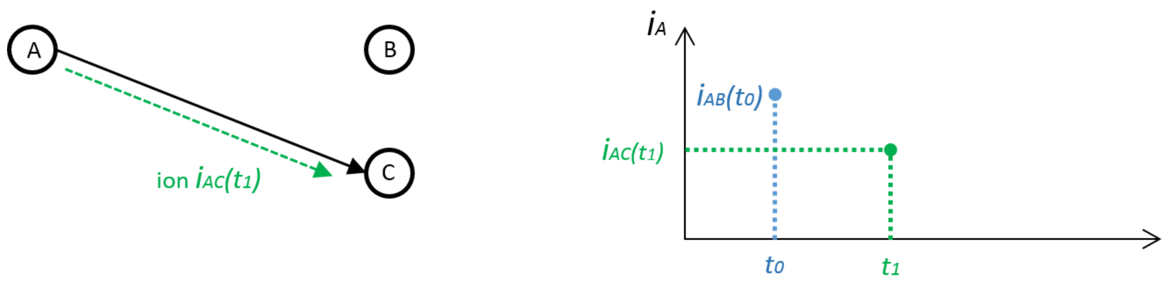
In general, we can say that:

$$i(t_1) = \sum i_k(t_1) = \sum i_k(t_0) \cdot \frac{1}{1 + e^{10\lambda(t_1 - t_0) - 5}}$$

where the series `i_k` is a history of all attention ions sent in the past.

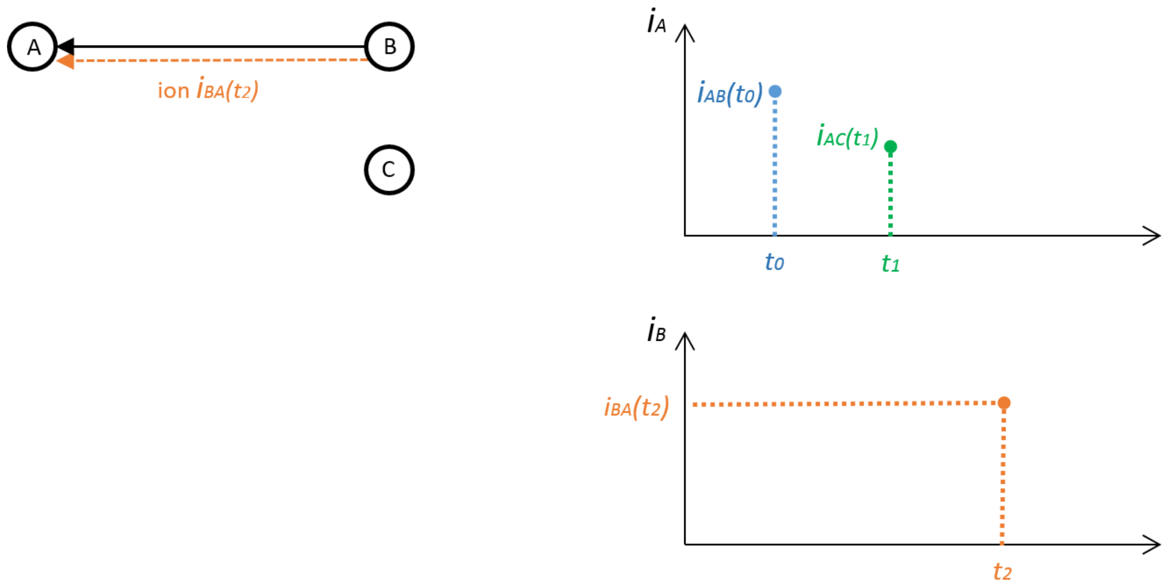
Calculating Attention Intensity

Let us return to the example from above where person A has fired an attention ion to B. Now suppose that at some time later person A sends an attention ion `i_AC` to person C.



Person A sends an attention ion to person C

And a little later, person B sends attention ion `i_BA` back to person A.



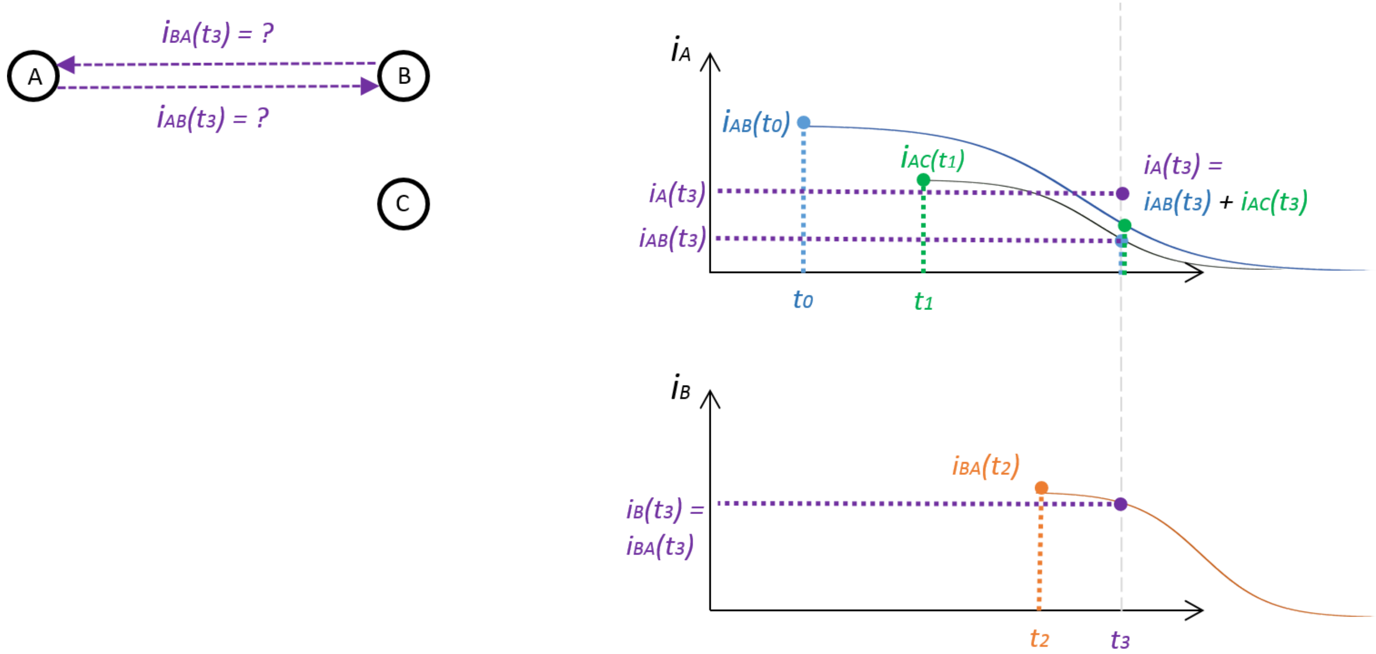
Person B sends an attention ion back to person A

Later at time `t_3`, we want to calculate the attention intensities `i_AB(t_3)` and `i_BA(t_3)`. To do so, we go through the history for A and B and a) find any ions that were sent from A to B (or vice versa) and b) sum the ion functions. In this example, there is only ion in the history in each case, so the sum has only one term:

$$
\begin{aligned}
i_{AB}(t_3) &= i_{AB}(t_0) \cdot \frac{1}{1 + e^{10\lambda(t_3 - t_0) - 5}} \\
i_{BA}(t_3) &= i_{BA}(t_2) \cdot \frac{1}{1 + e^{10\lambda(t_3 - t_2) - 5}}
\end{aligned}
$$

This gives us the attention intensities being exchanged between B and A. Now, we also want the *total intensities being emitted by each A and B to **all** targets (not just each other)*. For this, we a) find *any* ion in the history emitted by A (or B) and b) sum the time-distributed values. In this example, for B this is a sum with only one term (again), but for A we end up adding the distributed intensities from two previous ions:

$$i_A(t_3) = i_{AB}(t_3)+i_{AC}(t_3) = i_{AB}(t_0)\cdot\frac{1}{1+e^{10\lambda(t_3-t_0)-5}}+i_{AC}(t_1)\cdot\frac{1}{1+e^{10\lambda(t_3-t_1)-5}}$$
$$i_B(t_3) = i_{BA}(t_3) = i_{BA}(t_2)\cdot\frac{1}{1+e^{10\lambda(t_3-t_2)-5}}$$



Calculating attention intensities as sum of ion functions

Conservation of Attention Intensity

At this stage, let us pause to note that there is (as yet) no upper-bound to attention intensity here - if the number of ions in the history is large and if they are recent, the value of `i` at any given time can be increased to an arbitrarily high value. This is because we have not yet applied the constraint of *conservation of attention intensity*.

This constraint is copied directly from the gaze model as above. Let us define `i_MAX` as the maximum possible attention intensity from any single node, and without any loss of generality we set to `1`. We then introduce the normalizing factor `K(t)` to multiply every attention intensity:

$$\widehat{i_k(t)} = K(t)\cdot i_k(t)$$

where `i^` is the normalized value of attention intensity. `K(t)` is such that:

$$K(t)\cdot\sum i_k(t) \geq 1$$

and `K(t)` can be explicitly calculated as:

$$if \sum i_k(t) \leq 1\, then\, K(t) = 1\ else\ K(t) = \frac{1}{\sum i_k(t)}$$

Note that the calculation of `K` requires all intensities to all targets calculated and summed, even if it is being calculated only to normalize one particular value of `i` for only one target. For example, if we wish to calculate the normalized value of `i^_AB`, we need to calculate both `i_AB` and `i_AC` to calculate the value of `K`, which is then used to get `i^_AB`.

This point is worth repeating as an important result:

the calculation of any (normalized) ion intensity from A to B requires knowledge of all emitted ions from A to all targets in the network

Since we are always interested in the normalized values for intensities, *we will now drop the notation for `i^` and proceed with the convention that all attention intensities `i` are henceforth referring to the normalized value*.

Also note that `i(t)` satisfies the constraint `0 <= i(t) <= 1` at all times.

Calculating Mutual Attention

It is now possible to calculate the skew and affordance between A and B at any time `t` as a function of the (normalized) intensity values `i_AB` and `i_BA`:

$$S(t) = \sqrt{|i_{AB}^2 - i_{BA}^2|}$$

$$A(t) = \sqrt{\frac{i_{AB}^2 + i_{BA}^2}{2}}$$

Giving the intima as a function of time:

$$\mathbf{M}(t) = [S(t), A(t)]$$

Analysis and Notes

Composite Ion Function

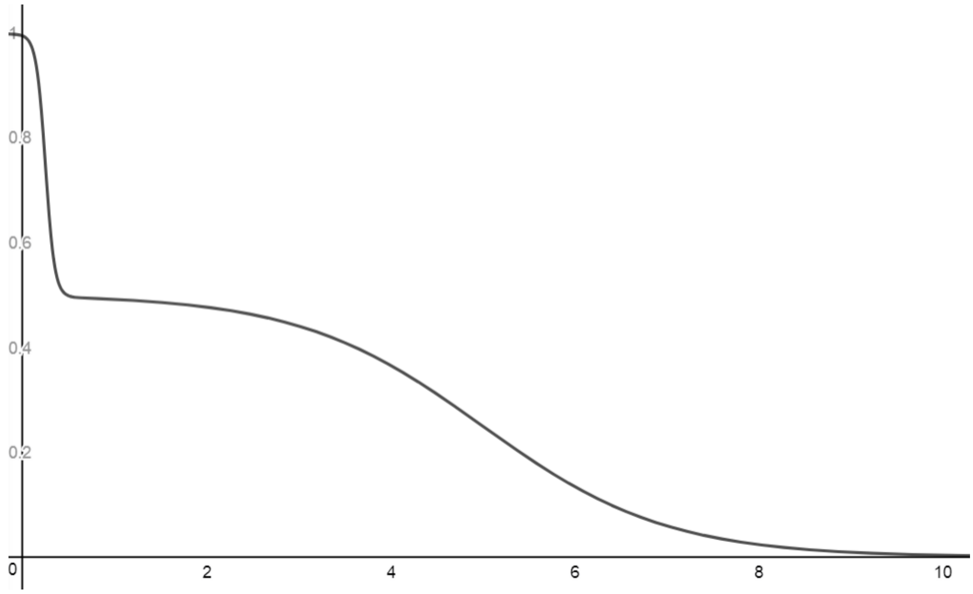
The ion function may be enhanced by making it a sum of two sigmoidal functions:

$$i(t_1) = i_0 \cdot \frac{W_0}{1 + e^{10\lambda_0 t - 5}} + i_0 \cdot \frac{W_1}{1 + e^{10\lambda_1 t - 5}}$$

where we have parameters `lamba_0` and and `lambda_1` and we choose the factors `W` such that:

$$W_0 + W_1 = 1$$

Now we choose `lambda_0` to be large (fast decay) and `lambda_1` to be small (slow decay). This gives us a composite curve something like this:



Composite ion function

This is a useful enhancement because it adds a new behaviour to the ion function that allows it to model a new quality of human interactions - whilst the fast decay continues to model the short-term attention signalling used in live conversation (gaze), the slow decay models the longer term attention of an ongoing collaboration or relationship (mediated by memory).

Natural Trajectories in Mutual Attention Space

We now consider some *natural trajectories* through time in mutual attention space. Natural trajectories are canonical paths in time for mutual communication/attention. These are empirical and describe human social dynamics that are typically observed in the online world.

[Note that these trajectories are idealized and continuous for purposes of illustration. In the real implementation of Intima, the discrete nature of signalling with ions will create discontinuous trajectories that "jump" around.]

Firstly, if A and B have *never* exchanged attention at all, then `A = 0` and `S = 0`. This is the starting point for a completely new relationship between A and B and is known as the *zero mutual attention* point.

On the other hand, if A and B are engaged in maximum intensity mutual attention, then `A = 1` and `S = 0` and this is the *ideal mutual attention* point.

Lastly, if A is directing maximum intensity of attention at B with a complete lack of response (or vice versa) then `A = 1` and `S = 1` and this is known as the *one-way attention* point.

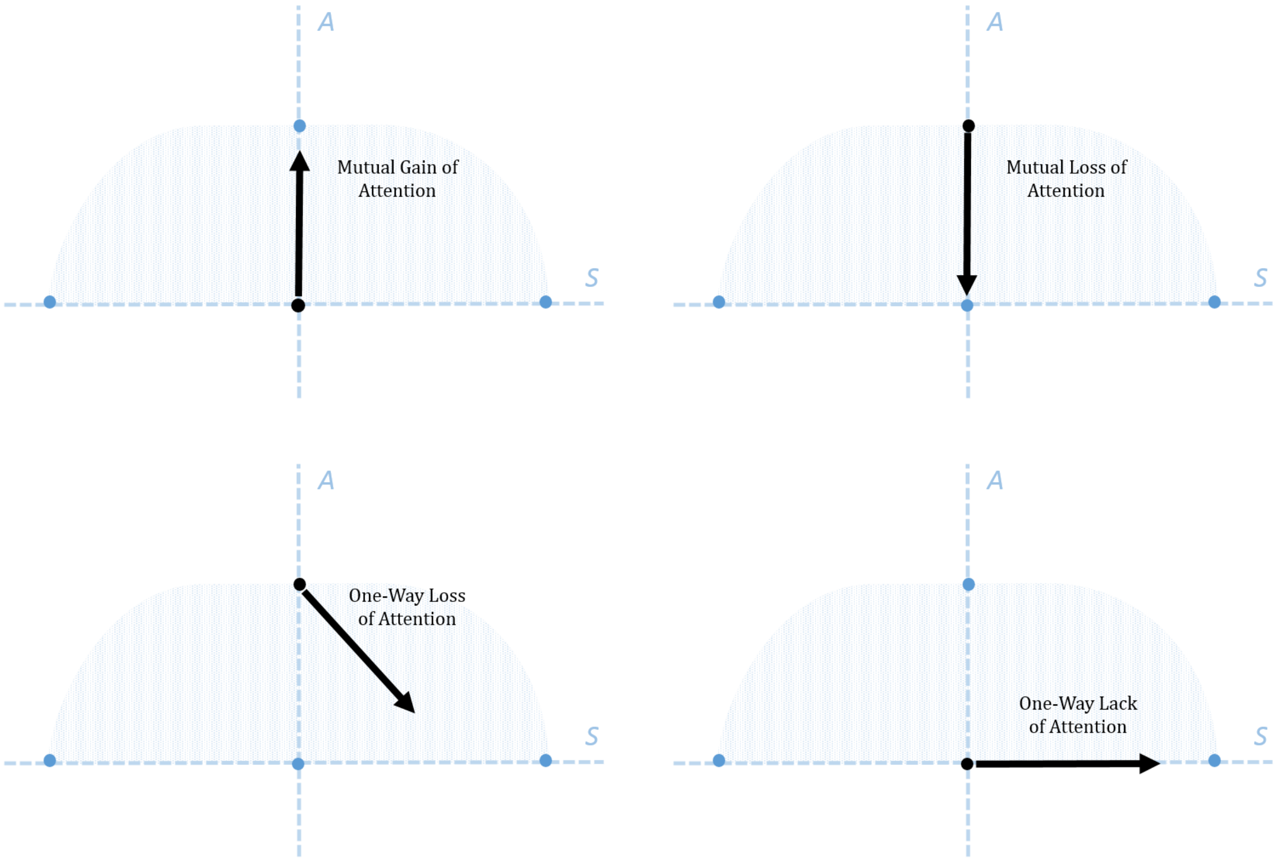
Now we can consider four different natural trajectories:

1. *mutual gain of attention* occurs when A and B interact for the first time and steadily exchange growing intensities of attention in a reciprocal manner. Example is a new friendship or business partnership.

2. *mutual loss of attention* occurs when A and B allow a high-M relationship to drift back toward zero mutual attention by mutually withdrawing attention from each other. Example is "losing touch" with friends or colleagues.

3. *one-way loss of attention* occurs when one person A allows a relationship to degrade toward one-way attention by withdrawing reciprocity i.e. ignoring the other person B. Example is "going cold" on a sales relationship.

4. *one-way lack of attention* occurs when when A and B interact for the first time but A completely fails to respond to attention from B. Example is ignoring a persistent spammer.



Natural trajectories in time through mutual attention space

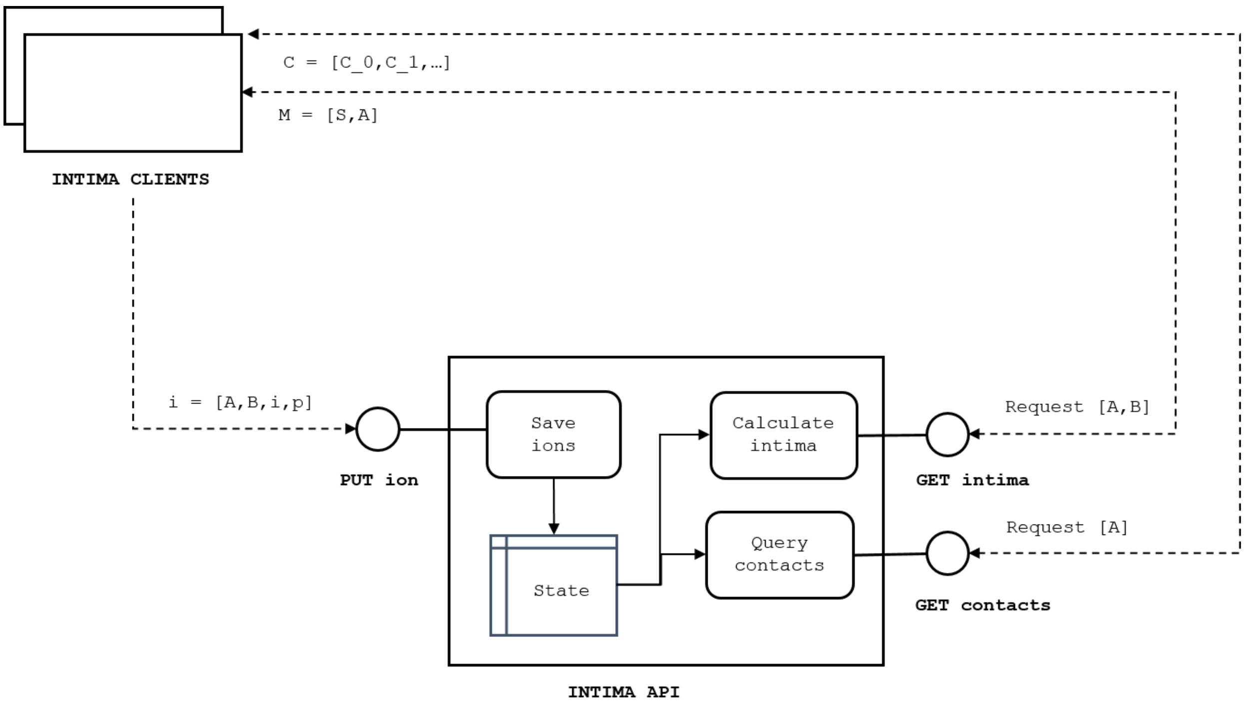
Implementation

In this section, we document the implementation of intima and how it is introduced into the existing networks.

Architecture

Intima is implemented as a centralized REST API that interacts with client over HTTPS. Clients make requests to the API which provides three interface endpoints:

1. **PUT an ion** - receives an ion `i_AB` and adds to internal state with current time
2. **GET intima** - for requested nodes A and B, returns intima `M_AB` calculated at current time
3. **GET contacts** - for requested A, returns nodes `[C]` which have non-zero intima values `M_AC`



Intima system architecture

[TODO]

Simulations

[TODO]

Applications

In this section we introduce some applications of intima.

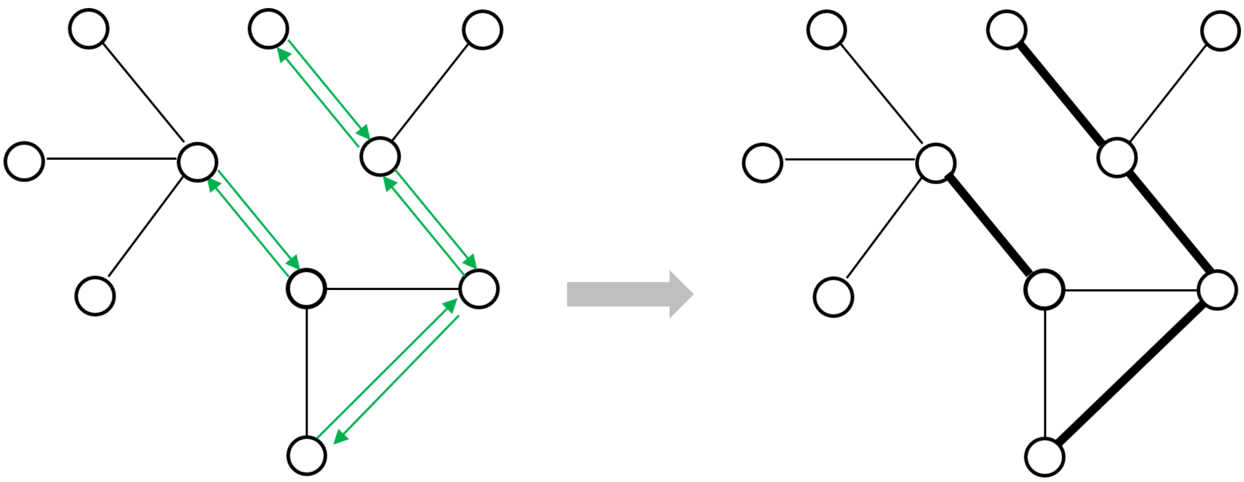
Improved QoS in Mobile Telecommunication Networks

Quality of Service (QoS) for mobile networks relies on effective resource reservation/allocation. New allocation algorithms that are functions of intima (affordance and skew) will reward network edges that have mutual attention with increased accessibility, audio quality and Grade of Service (GoS).

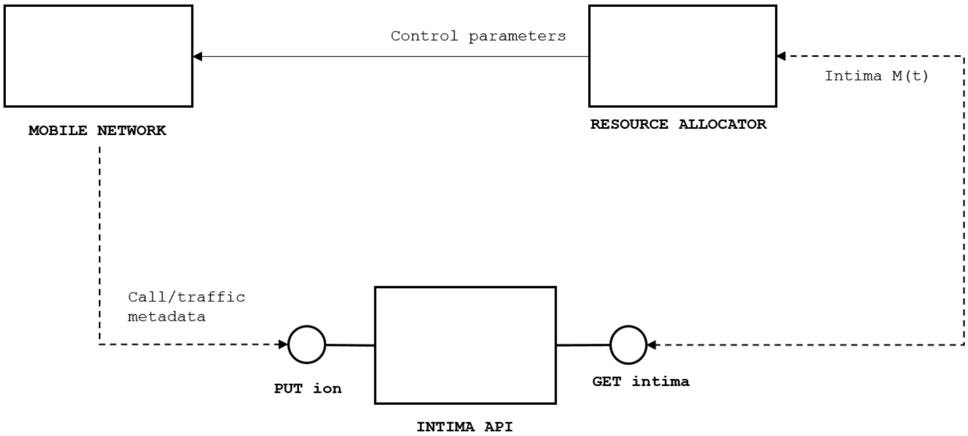
This can be implemented at more than one time-scale. For example, network edges with sudden spikes in affordance (real-time conversation) can be rewarded with bursts of resource allocation; edges with a longer-term history of mutual attention with more static resource reservations.

This "edge-centric" intima algorithm is more responsive and dynamic than the conventional "node-centric" (per-user) allocation strategies.

[TODO: simulation/PoC to validate this intuition]



QoS allocation with intima - edges with mutual attention get more resources



QoS allocation with intima - system design

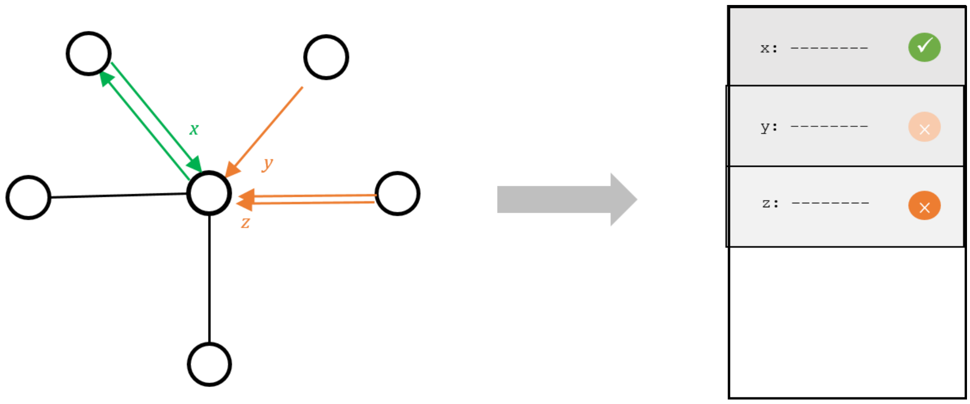
Improved Dynamic Bandwith Allocation

Dynamic Bandwidth Allocation techniques can be improved by the use of intima, rewarding network edges that have mutual attention with increased bandwidth allocations. This privileges mutual communication over one-way traffic, a property that will be useful in many distributed network situations e.g. instant messaging networks.

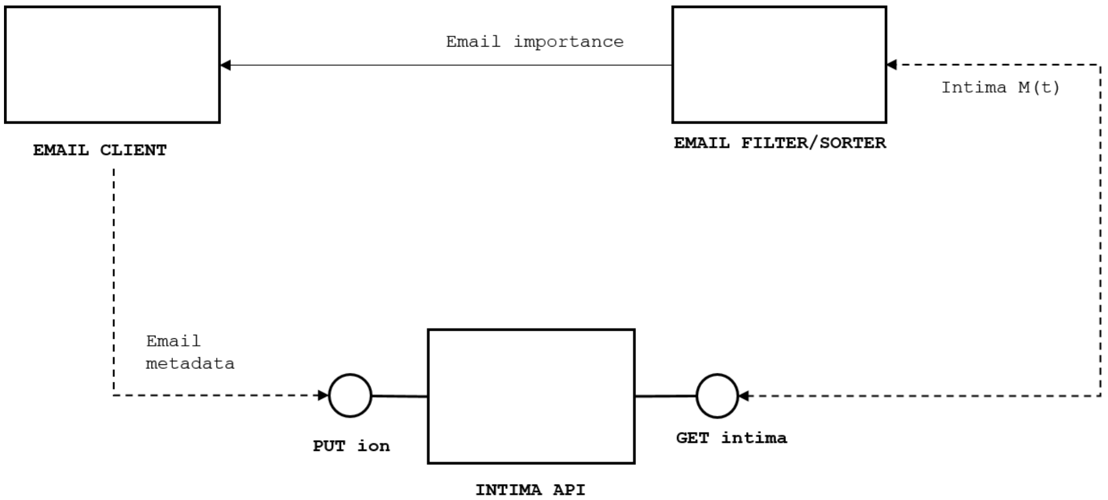
Embedded, Continuous Spam Filtering for Email Clients

Current schemes for spam filtering work on a discrete 1/0 scheme, scoring incoming emails on a number of continuous parameters but using a thresholding scheme to classify emails as either SPAM or NOT SPAM.

Next generation spam filtering will use intima, looking for incoming emails on edges with high skew values (indicating a history of one-way/unreciprocated attention) and rank/sort emails accordingly. Also, edges with history of high mutual attention get promoted to top of the list, as they are likely to be trusted contacts.



Email ranking with intima - one-way attention means spam, mutual attention means important

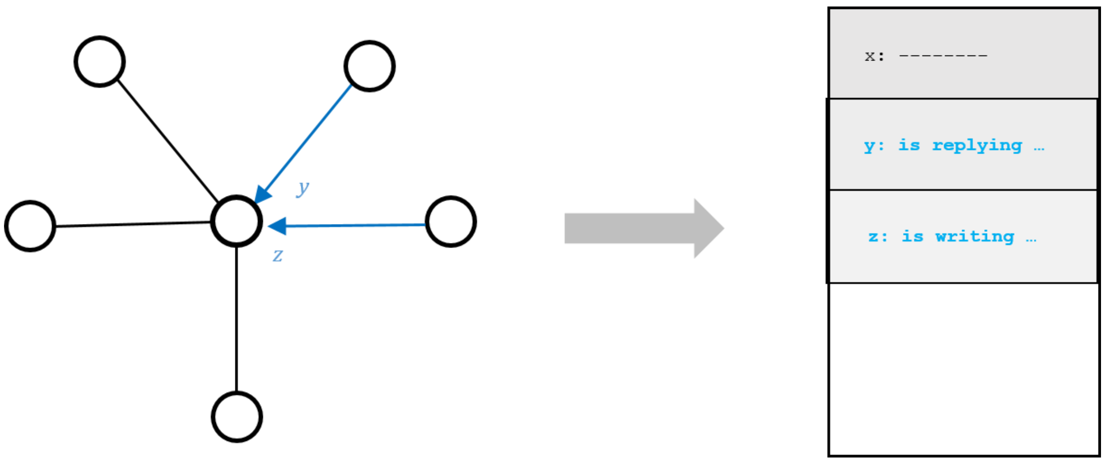


Email ranking with intima - system design

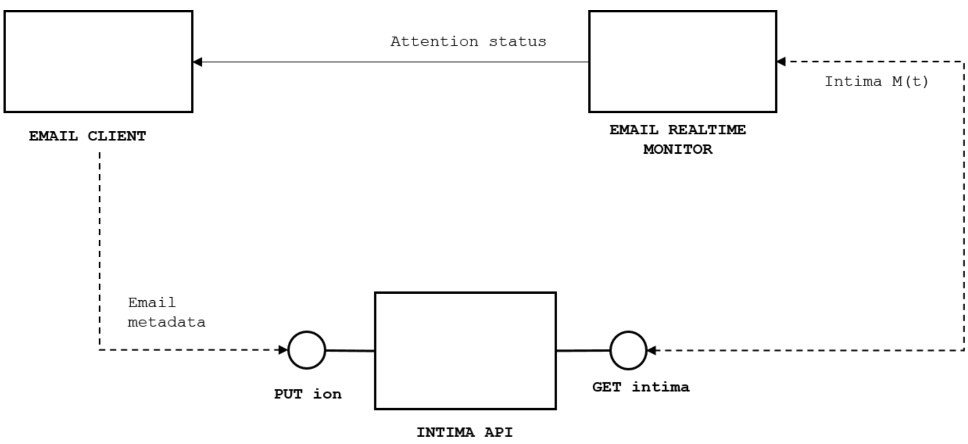
[TODO]

Augmented Email Network with Real-Time Signalling

Email lacks the real-time feedback mechanisms which are now standard for other messaging platforms such as Skype and WhatsApp (e.g. "XXX is typing a message", "XXX has seen your message" etc.). Intima provides an easy way to retrofit an email network with feedback, as sudden spikes in affordance indicate attention on a network edge. Additionally, intima can be enhanced to return directional attention intensities in addition to skew and affordance for even more fine-grained feedback.



Email attention signalling - augmenting email UX with attention signals



Email attention signalling - system design

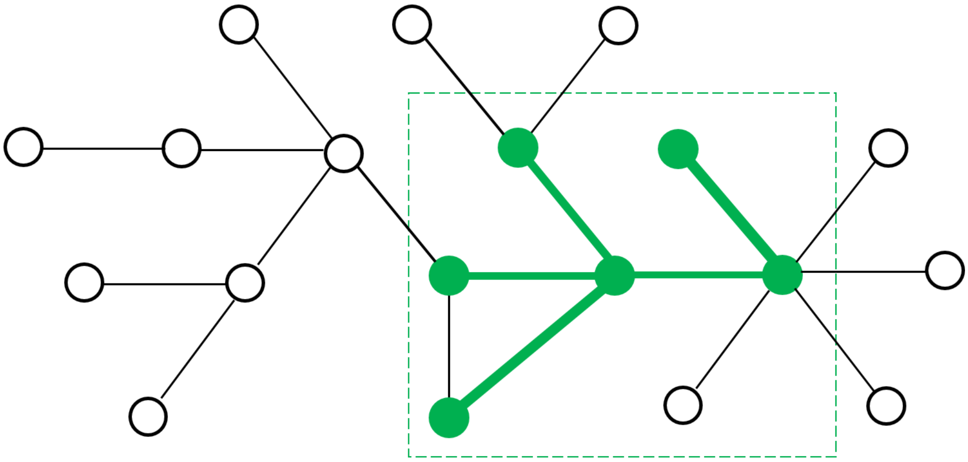
[TODO]

Standalone Attention Signalling for Mobile Clients

[TODO]

Automated Grouping of Mutually Attentive People

An interesting application of intima is to apply a clustering algorithm to the intima network to find contiguous sub-networks that are connected with edges of high mutual attention (affordance). These sub-networks change over time, but at any time this provides a novel way to query for "instant teams"/"instant social networks".



Clustering to find sub-networks of high mutual attention

Automated Sales Feedback Loops

Automate feedback loops on which sales are "hot" vs "going cold". Focus in on "hot" sales leads. Track relationship strength. [TODO]