A data specification for efficient archival of ecological networks

Timothée Poisot

Feb. 2013

This is a *working document* describing `mangal`, a set of `JSON` objects templates to encode ecological networks of virtually any complexity. There are plans to host a pilot database.

The need for a data specification

Ecological networks enable ecologists to accommodate the complexity of natural communities, and to discover mechanisms contributing to their persistence, stability, resilience, and functioning.

- we need *a lot* of information to make inference and train predictive models
- networks are *hard* to get in nature
- need to compare several datasets to start working on the biogeography of web

A short paragraph describing the type of data we need and relating them to outstanding questions, ref from the recent literature. Space, evolution, gradients, prediction.

The elements of a network

Basic idea is to have *layers* in the data specification, each corresponding to one layer of organization in the networks. What I see currently is

- nodes
- edges
- networks

This will require some *glue* objects, but it should be simple enough.

The properties of nodes

- scale of organization
- biodiversity descriptors

The properties of edges

The properties of networks

- sample location
- sample date
- methodology
- associated references

Proposed implementation

This will mostly describe several data templates for each of the scales in the previous section.

```
{
    "ID": "MyDB_I_00001",
    "from": "MyDB_P_00001",
    "to": "MyDB_P_00002",
    "info": {
        "binary": true,
        "directional": false,
        "strength": 1,
        },
    "type": "predation",
    "notes": "none"
}
```

Conclusion and future directions