

УДК 631.811.98

ОБРАБОТКА ПАТЕНТНОЙ ИНФОРМАЦИИ ПЕСТИЦИДНОАКТИВНЫХ СОЕДИНЕНИЙ С РОСТОРЕГУЛИРУЕМЫМ ВИДОМ АКТИВНОСТИ

Шуваев А.В.

к.х.н., доцент

shuvaev53@mail.ru

ФГБОУ ВО Сибирский государственный университет путей сообщения,
г. Новосибирск, Россия

Аннотация. Проведен анализ патентных данных, содержащих информацию о пестицидноактивных химических соединениях с росторегулируемым видом активности. Разработаны наиболее типичные форматы, с помощью которых произведена обработка результатов биологических тест-испытаний с последующим вводом информации в базу данных.

Ключевые слова: биоактивные соединения, химическая структура, обработка патентной биологической информации, база данных, генерация биологической информации.

В ряде предыдущих работ [1, 7 – 10] были изложены методологические основы построения базы данных биологически активных химических соединений. Вначале осуществляется предварительная обработка исходной патентной информации, которая представляется посредством отдельных текстов-файлов. Затем информация записывается на магнитные носители в виде нескольких блоков данных. Блок патентных характеристик – основных сведений о литературном источнике, содержится в файле "Patent". Блок химической структурной информации – набор файлов "Atom", "Radical", "Formula", "Replace" [7, 8]. Блок физико-химической и биологической информации – файлы "Format", "Value" [1, 9].

Как уже отмечалось ранее [9, 10], первоначально наибольшее внимание было уделено химической части проблемы, в результате чего разработан алгоритм обработки и ввода в компьютерную память химической структурной информации. Решение биологической части проблемы осуществляется путем формирования двух файлов: "Format" – перечень названий наиболее важных характеристик; "Value" – экспериментальные данные о биологической активности для конкретного химического соединения и испытуемого тест-организма. При этом формирование файла "Value" и заполнение его позиций должно однозначно соответствовать позициям формата (форматов) файла "Format". Таким образом, основным звеном в решении проблемы обработки биологической информации является создание формата. Попытки создания единого универсального формата не привели к успеху, поэтому было решено проводить обработку биологической информации поэтапно, отдельно в зависимости от принадлежности к определенному классу пестицидов – гербициды, фунгициды, регуляторы роста растений, инсектициды и др., и, по мере возможности, стремиться к созданию наиболее общего формата.

В работах [9, 10] подробно изложены результаты обработки биологических данных для ряда химических соединений, проявляющих гербицидную и фунгицидную активность. В данной работе исследуются биологические данные соединений другого вида активности – регуляторы роста растений.

Процедура обработки биологической информации. Файл "Format"

Любой формат, соответствующий биологическим данным, должен

состоять из двух характерных частей: вспомогательной и основной. Вспомогательная часть отражает информацию общего характера: номер формата (цифра); номер магнитной ленты или часть жесткого диска (ML) – условный номер одного или нескольких совместно обрабатываемых патентов; структурный интервал, охватываемый патентом (structure interval); химический класс соединений (chemical class); вид активности; название теста (test); описание метода (application mode); единицы биологической активности (numerical scale); оценочный критерий (rating term); наименование растения (test-organism); фаза обработки (test regime); среда обитания (growth medium).

Основная часть формата связана с размещением экспериментальных данных – физико-химической и биологической информации: структурный номер испытуемого соединения (structure number); температура плавления твердых веществ (°C) (m.p.(grad C)); для жидких и индивидуальных соединений – температура кипения при определенном значении давления (b.p. (grad C) / p (torr)); показатель преломления при определенной температуре (ND / (grad C)); спектральные характеристики (IR (1 / SM)); доза (Rate (KG / HA)); концентрация (concentration (%)); биологическая активность (biology activity).

Обработкой данных патентов, содержащих информацию для различных химических соединений, обладающих росторегулируемым видом активности по отношению к определенным видам растений, было изготовлено и размещено в базе данных около 60 форматов, в файле "Format" им присвоены следующие номера: 71–107, 109, 111, 113–132. В качестве примера ниже приведены несколько форматов из этих диапазонов, использованы данные работ [2 – 6].

*Format:

74 & ML = & structure interval = & numerical scale: рост соответствует контролю (%) = & полное прекращение роста (%) = & задержка роста (%) = & стимуляция роста (%) = & test name = & test-organism = & Growth medium = & test regime = & application mode = & test-organism stage at treatment = & rating term = & structure number = & rate (KG/HA) = & концентрация вещества (%) = & Biology activity = *

78 & ML = & structure interval = & numerical scale: зрелые плоды отсутствуют(%) = & все плоды зрелые (%) = & test name = & test-organism = & Growth medium = & test regime = &

application mode = & test-organism stage at treatment = & rating term = & structure number = & rate (KG/HA) = & концентрация вещества (%) = & Biology activity = *

82 & ML = & structure interval = & numerical scale: длина стебля = & урожайность = & test name = & test-organism = & Growth medium = & test regime = & application mode = & test-organism stage at treatment = & rating term = & structure number = & rate (KG/HA) = & концентрация вещества (вес. %) = & Biology activity = *

97 & ML = & structure interval = & высота растений = & число кистей = & число цветков = & урожай = & test name = & test-organism = & Growth medium = & test regime = & application mode = & test-organism stage at treatment = & rating term = & structure number = & rate (KG/HA) = & концентрация вещества (вес, %) = & Biology activity = *

109 & ML = & structure interval = & вид активности = & химический класс = & оценочный критерий = & название теста = & вид метода = & описание метода = & номер метода = & фаза обработки = & тест-объект = & structure number = & Rate = & ед. дозы = & max. act. = & Biology activity = *

Файл "Value"

Процедурные действия, связанные с формированием файла "Value", изложены в работе [9]. Применительно к регуляторам роста растений следует отметить, что в этом файле хранится информация о биологической активности химических соединений по отношению к сельскохозяйственным культурам. Информация имеет цифровой или текстовый вид, причем порядок ее расположения строго соответствует последовательности позиций определенного формата. В том случае, когда обрабатываются данные одного патента, вспомогательная часть формата будет заполняться информацией однократно, так как носит постоянный характер для всего структурного интервала. Формирование основной части формата связано с размещением результатов испытаний действий биологической активности соединений на определенных растительных культурах. Иногда в процессе записи информации в этой части формата часть позиций может оказаться пустой, например, в патенте не имеется соответствующих данных, или, наоборот, возникает необходимость расширить ряд позиций в том случае, когда для одного и того же соединения приводится несколько результатов исследований при вариации условий испытаний. Для этих случаев можно использовать команды перемещения вправо или влево на

несколько позиций формата: операция: # + n обозначает перенос на n позиций в прямом направлении; операция: # – m – возврат влево на (m – 1) позицию.

Химические соединения, относящиеся к разнообразным классам соединений, исследуются разными авторами с позиций выяснения характера влияний на рост растений. В большинстве работ методика испытаний сводится к следующему: семена или растения в определенной стадии развития обрабатываются препаратом в виде растворов, эмульсий, суспензий, дустов в различных дозах. Через определенный период времени оценивают результаты с контрольными данными – растения без обработки. При этом регулирующий рост растений активность химических соединений может быть двояка: увеличивающая, обычно используемая для прибавки веса урожая, и замедляющая – укорачивание стебля растения для устойчивости или уменьшения периода цветения с целью ускорения созревания плодов. Оценочным критерием является разница высоты роста в сантиметрах или процентный показатель. В качестве тест-объектов используются различные культуры: горох, свекла, ячмень, пшеница, кукуруза, огурцы, томаты, бобы, фасоль, яблони, рис, хлопок, цветы и др.

Для наглядности ниже приведена часть файла "Value", в которой отражена информация о росторегулирующем виде активности для определенных химических соединений, использованы данные работ [2 – 6].

*Value:

3 - 74 & 4891 - 4946 & 0 & - 100 & - 45 & + 34 & пример 4 & ячмень & - & - & Post-emergence & 2 листа & по достижении в контроле высоты растений 60 см & 4919 & 0.05 - 10 & 0.05 & - 65 & # - 4 & 4918 & 0.05 - 10 & 0.05 & - 35 & # - 4 & 4908 & 0.05 - 10 & 0.05 & - 65 & # - 4 & 4892 & 0.05 - 10 & 0.05 & - 75 & # - 4 & 4910 & 0.05 - 10 & 0.05 & - 30 & # - 4 & 4897 & 0.05 - 10 & 0.05 & - 75 & # - 4 & 4915 & 0.05 - 10 & 0.05 & - 40 & # - 4 & 4906 & 0.05 - 10 & 0.05 & - 40 & # - 4 & 4925 & 0.05 - 10 & 0.05 & - 35 & # - 4 & 4898 & 0.05 - 10 & 0.05 & - 35 & # - 4 & 4927 & 0.05 - 10 & 0.05 & - 35 & # - 4 & 4935 & 0.05 - 10 & 0.05 & - 85 & # - 4 & 4941 & 0.05 - 10 & 0.05 & - 35 & # - 4 & 4945 & 0.05 - 10 & 0.05 & - 80 & # - 4 & 4939 & 0.05 - 10 & 0.05 & - 60 & # - 4 & 4901 & 0.05 - 10 & 0.05 & - 40 ...

3 - 78 & 5105 - 5124 & 0 & 100 & пример 5 & томаты && in a greenhouse & Post-emergence & 30 % зрелых плодов & через 2 дня после обработки & 5106 && 0.1 & 53 & # - 5 & через 2 дня после обработки & 5106 && контроль & 37 & # - 5 & через 4 дня после

обработки & 5106 && 0.1 & 65 & # - 5 & через 4 дня после обработки & 5106 && контроль & 40 & # - 5 & через 7 дней после обработки & 5106 && 0.1 & 86 & # - 5 & через 7 дней после обработки & 5106 && контроль & 44 ...

3 - 82 & 5185 - 5186 & % к контролю && пример 1 & wheat & soil + 1 г NH₄NO₃ + 1,2 г СаНРО₄ + 0,8 г К₂SO₄ + 0,5 г MgSO₄ · 7 Н₂О в сосудах Митчерлиха & ежедневный полив 50 % от максимальной влагоемкости почвы, в теплице & полив & 1 дневные всходы & к моменту полного созревания & 5185 & 2 & 1 & 73.1 & # - 11 & пример 2 & wheat & soil & в поле & Post-emergence & всходы & к моменту полного созревания & 5185 & 2 & 1 & 93 & # - 12 & % к контролю & пример 2 & wheat & soil & в поле & Post-emergence & всходы & к моменту полного созревания & 5185 & 2 & 1 & 93 & # - 12 & % к контролю & пример 2 & wheat & soil & в поле & Post-emergence & всходы & к моменту полного созревания & 5185 & 2 & 1 & 116.0 & # - 11 & пример 3 & Barley & soil & в поле & Post-emergence & всходы & к моменту полного созревания & 5185 & 1 & 1 & 98 ...

3 - 97 & 5231 - 5231 & см & # + 3 & пример 3 & томат грунтовый грибоедовский & soil & в полевых условиях & Pre-emergence, замачивание семян & seeds & к моменту уборки урожая & 5231 & 0.001 & 121.3 & # - 2 & 0.0 & 111.6 & # - 14 & штук & # + 2 & пример 3 & томат грунтовый грибоедовский & soil & в полевых условиях & Pre-emergence, замачивание семян & seeds & к моменту уборки урожая & 5271 && 0.001 & 17.0 & # - 2 & 0.0 & 14.1 & # - 13 & штук && пример 3 & томат грунтовый грибоедовский & 1180 & soil & в полевых условиях & Pre-emergence, замачивание семян & seeds & к моменту уборки урожая & 5231 & 0.001 & 139 & # - 2 & 0.0 & 130 & # - 12 & ц/га & пример 3 & томат грунтовый грибоедовский & 1180 & soil & в полевых условиях & Pre-emergence, замачивание семян & seeds & к моменту уборки урожая & 5231 && 0.001 & 131.8 & # - 2 & 0.0 & 111.7 ...

2 - 109 & 9360 - 9403 & PPP & соли алкиламмония & высота роста, см & высота роста, % & пример 1 & тепл. & растение после всхода обрабатывают препаратом ДВ (20 % вес. %).
Оценку проводят через 31 день после обработки & 1 & после всходов & соя сорта SRF 400 & 9364 & 1.5 & мг ДВ на сосуд & % & 25.3 / 100 & 15.0 / 59.3 & # - 5 & 6.0 & мг ДВ на сосуд & % & 25.3 / 100 & 13.5 / 53.4 & # - 6 & 9369 & 1.5 & мг ДВ на сосуд & % & 25.3 / 100 & 21.5 / 85.0 & # - 5 & 6.0 & мг ДВ на сосуд & % & 25.3 / 100 & 15.0 / 59.3 & # - 6 & 9370 & 1.5 & мг ДВ на сосуд & % & 25.3 / 100 & 16.0 / 63.2 & # - 5 & 6.0 & мг ДВ на сосуд & % & 25.3 / 100 & 13.5 / 53.4 & # - 6 & 9371 & 1.5 & мг ДВ на сосуд & % & 25.3 / 100 & 17.0 / 67.2 & # - 5 & 6.0 & мг ДВ на сосуд & % & 25.3 / 100 & 15.0 / 59.3 & # - 6 & 9376 & 1.5 & мг ДВ на сосуд & % & 25.3 / 100 & 18.0 / 71.2 & # - 5 & 6.0 & мг ДВ на сосуд & % & 25.3 / 100 & 16.5 / 65.2 & # - 6 & 9377 & 1.5 & мг ДВ на сосуд & % & 25.3 / 100 & 16.5 / 65.2 & # - 5 & 6.0 & мг ДВ на сосуд & % & 25.3 / 100 & 14.5 / 57.3 & # - 6 & 9378 & 1.5 & мг ДВ на сосуд & % & 25.3 / 100

& 15.0 / 59.3 & # - 5 & 6.0 & мг ДВ на сосуд & % & 25.3 / 100 & 14.0 / 55.3 ...

В дальнейшем в режиме "Activity" происходит объединение содержимого файлов "Format" и "Value" для конкретного химического соединения и информация в сгенерированном виде выводится на экран монитора или на печать. Ниже продемонстрирована эта процедура на ряде примеров.

*Activity:

ML = 3 & structure interval = 4891 - 4946 & numerical scale: рост соответствует контролю (%) = 0 & полное прекращение роста (%) = - 100 & задержка роста (%) = - 45 & стимуляция роста (%) = + 34 & test name = пример 4 & test-organism = ячмень & Growth medium = - & test regime = - & application mode = post-emergence & test-organism stage at treatment = 2 листа & rating term = по достижении в контроле высоты растений 60 см & structure number = 4919 & rate (KG/HA) = 0.05 - 10 & концентрация вещества (%) = 0.05 & biology activity = - 65 *

*Activity:

ML = 3 & structure interval = 5105 - 5124 & numerical scale: зрелые плоды отсутствуют (%) = 0 & Все плоды зрелые (%) = 100 & test name = пример 5 & test-organism = томаты & Growth medium = - & test regime = in greenhouse & application mode = post-emergence & test-organism stage at treatment = 30 % зрелых плодов & rating term = через 2 дня после обработки & structure number = 5106 & rate (KG/HA) = & концентрация вещества (%) = 0.1 & biology activity = 53 & rating term = через 2 дня после обработки & structure number = 5106 & rate (KG/HA) = & концентрация вещества (%) = контроль & biology activity = 37 & rating term = через 4 дня после обработки & structure number = 5106 & rate (KG/HA) = & концентрация вещества (%) = контроль & biology activity = 40 & rating term = через 7 дней после обработки & structure number = 5106 & rate (KG/HA) = & концентрация вещества (%) = 0.1 & biology activity = 86 & rating term = через 7 дней после обработки & structure number = 5106 & rate (KG/HA) = & концентрация вещества (%) = контроль & biology activity = 44 *

*Activity:

ML = 3 & structure interval = 5231 - 5231 & высота растений = см & число кистей = & число цветков = & урожай = & test name = томат грунтовый грибоедовский & Growth medium = soil & test regime = в полевых условиях & application mode = pre-emergence, замачивание семян & test-organism stage at treatment = seeds & rating term = к моменту уборки урожая & structure number = 5231 & rate (KG/HA) = & концентрация вещества (вес, %) = 0,001 & biology activity = 121.3 & концентрация вещества (вес, %) = 0.0 & biology activity = 111.6 *

*Activity:

ML = 2 & structure interval = 9360 - 9403 & вид активности = PPP & химический класс

= соли алкил аммония & оценочный критерий = высота роста, см / высота роста, % & название теста = пример 1 & вид метода = тепл. & описание метода = растения после всхода обрабатывают препаратом ДВ (20 вес %). Оценку производят через 31 день после обработки & номер метода = 1 & фаза обработки = после всходов & тест-объект = соя сорта SRF 400 & structure number = 9364 & rate = 1.5 & ед. дозы = мг ДВ на сосуд & max act. = % & min act. = 25.3 / 100 & biology activity = 15.0 / 59.3 & Rate = 6.0 & ед. дозы = мг ДВ на сосуд & max act. = % & min act. = 25.3 / 100 & biology activity = 13.5 / 53.4 *

Таким образом, результаты проделанного анализа большого числа патентных данных биологической информации пестицидноактивных соединений с росторегулируемым видом активности позволяют сделать следующие выводы. Сопоставление методик исследования разных авторов по выявлению эффективности действия определенных химических соединений в качестве регуляторов роста растений показало на наличие различий в части условий проводимого эксперимента. В основном они сводятся к следующему: в одних работах испытания осуществляли в полевых, а в других – в лабораторных условиях; изучаются различные виды растений; разные стадии обработки – предпосевная и послевсходовая; действия на разных стадиях всхожести растений – семена, проростки, взрослое растение; изучаются воздействия на разные части растений – корневая система, стебли, листья, плоды; различные концентрации обрабатываемых растворов; различия в способах обработки – замачивание, опрыскивание, опыление и др. Для того, чтобы не потерять специфику этой информации при обработке патентных данных с последующим вводом информации в базу данных, потребовалась разработка ряда типичных форматов. Их можно использовать в дальнейшем для решения локальных задач поиска необходимой информации по выбору определенного типа препарата при конкретных условиях или при обратной постановки вопроса: путем генерации информации базы данных в виде "структура – вид биологической активности – условия эксперимента – результаты воздействия".

В то же время следует отметить, что общим объединяющим звеном этих работ является форма представления конечных результатов в виде критерия оценки воздействия препарата – относительного изменения роста растения по

отношению к контролю (без действия препарата) в процентах. Это особенно важно для последующей статистической обработки большого массива исходных данных с целью поиска корреляционных зависимостей вида: "структурное звено – характеристика биологической активности" и выявления наиболее статистически значимых структурных фрагментов молекул соединений по тому или другому виду биологической активности.

Список литературы

1. Гаврилов Д.А., Шуваев А.В. Обработка патентной информации и создание базы данных соединений с росторегулируемой активностью: Химия и жизнь: сб. статей XVIII междунар. науч.-практ. конф./ Новосибир. гос. аграрн. ун-т. – Новосибирск, 2019. – С. 57 – 62.
2. Заутер Г., Цее Б., Бушман Б., Юнг И. Регулятор роста растений. Патент ФРГ, № 2913523.7. Заявл. 04.04.79: опубл. 26.03.80, 22 с.
3. Смирнов П.С., Базыльчик Н.М., Рябушкина Н.М. Регулятор роста растений. Патент СССР, № 782788. Заявл. 30.05.79: опубл. 30.11.80, 6 с.
4. Томас Р., Люрссен К. Средство для регулирования роста растений. Патент ФРГ № 2601376.3. Заявл. 13.01.77: опубл. 15.05.78, 18 с.
5. Хофер В., Люрссен К., Шмидт Р.Р. Регулятор роста растений. Патент ФРГ № 2110773.7. Заявл. 25.02.72: опубл. 30 04.76, 10 с.
6. Шиллинт Г., Бертнер Х., Отто Ш. Средство для регулирования роста растений. Патент ГДР, № 174690. Заявл. 15.11.74: опубл. 08.02.79, 8 с.
7. Шуваев А.В. Основы построения базы данных биологически активных химических соединений для решения экологических вопросов: Вопросы строительства и инженерного оборудования объектов железнодорожного транспорта. Материалы научно–практической конференции. – СГУПС. – 2017. – С. 147 – 157.
8. Шуваев А.В. Подготовка патентной химической структурной информации к вводу в базу данных биологически активных химических соединений // Экономика: экономика и сельское хозяйство. – 2018. – № 2 (26). – С. 5.
9. Шуваев А.В. Биологический аспект проблемы построения базы данных

биологически активных соединений // Наука и Образование. – 2019. – № 2. – С. 23.
10. Шуваев А.В. Обработка патентной информации пестицидно активных соединений с фунгицидным видом активности // Наука и Образование. – 2020. – № 1. – С. 124.

UDC 631.811.98

PROCESSING OF PESTICIDE-ACTIVE COMPOUNDS WITH A GROWTH-REGULATED ACTIVITY ON THE BASE OF PATENT INFORMATION

Shuvaev A.V.

PhD in Chemistry,

docent,

shuvaev53@mail.ru

Siberian State University of Railway Engineering,

Novosibirsk, Russia

Annotation. The analysis of patent data containing information on pesticide-active chemical compounds with a growth-regulated type of activity was carried out. The most typical formats have been developed, with the help of which the results of biological test-tests were processed and then entered into the database.

Keywords: bioactive compounds, chemical structure, biological processing of patent information, the database, the generation of biological information.